

TESIS - PM092315

**IDENTIFIKASI MAHASISWA YANG MEMPUNYAI
KECENDERUNGAN LULUS TIDAK TEPAT WAKTU
PADA PROGRAM STUDI MMT-ITS DENGAN
MENGUNAKAN ALGORITMA C4.5**

AMELIA HALIM
NRP 9112205304

DOSEN PEMBIMBING
Prof. Dr.Ir. Joko Lianto Buliali, M.Sc.

PROGRAM STUDI MAGISTER MANAJEMEN TEKNOLOGI
BIDANG KEAHLIAN MANAJEMEN TEKNOLOGI INFORMASI
PROGRAM PASCASARJANA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

TESIS - PM092315

IDENTIFICATION OF STUDENTS WHO HAVE GRADUATED NOT IN TIME ON MMT-ITS USING C4.5 CLASSIFICATION ALGORITHM

AMELIA HALIM
NRP 9112205304

DOSEN PEMBIMBING
Prof. Dr.Ir. Joko Lianto Buliali, M.Sc.

PROGRAM STUDI MAGISTER MANAJEMEN TEKNOLOGI
BIDANG KEAHLIAN MANAJEMEN TEKNOLOGI INFORMASI
PROGRAM PASCASARJANA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

IDENTIFIKASI MAHASISWA YANG MEMPUNYAI KECENDERUNGAN LULUS TIDAK TEPAT WAKTU PADA PROGRAM STUDI MMT-ITS DENGAN MENGGUNAKAN ALGORITMA C4.5

**Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Manajemen Teknologi (M.MT) di
Institut Teknologi Sepuluh Nopember**

**Oleh :
Amelia Halim
NRP 9112205304**

**Tanggal Ujian: 14 Januari 2015
Periode Wisuda: Maret 2015**

Disetujui oleh :

**1. Prof. Dr.Ir. Joko Lianto Buliali, M.Sc.
NIP. 196707271992031002**

(Dosen Pembimbing)

**2. Dr. Eng. Febriliyan Samopa, SKom, MKom
NIP. 197302191998021001**

(Dosen Penguji I)

**3. Erma Suryani, S.T., M.T., Ph. D.
NIP. 197004272005012001**

(Dosen Penguji II)

Direktur Program Pascasarjana

**Prof. Dr. Ir. Adi Soeprijanto, M.T.
NIP. 19640405 199002 1 001**

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena hanya berkat dan rahmat-Nya saja akhirnya Tesis ini dapat diselesaikan oleh penyusun. Tesis ini adalah tugas yang harus diselesaikan sebagai syarat kelulusan dari jurusan Manajemen Teknologi Informasi Magister Manajemen Teknologi Institut Teknologi Sepuluh November.

Penyusun mengucapkan terima kasih kepada Almamater Tercinta Magister Manajemen Teknologi Institut Teknologi Sepuluh November, atas segala prasarana yang telah disediakan dan sangat mendukung dalam mengikuti studi di Magister Manajemen Teknologi Institut Teknologi Sepuluh November. Atas segala bimbingan, dorongan, motivasi, yang telah diberikan, penyusun juga mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Tuhan Yang Maha Esa atas segala karunia-Nya.
2. Yang Terhormat Bapak Prof. Dr.Ir. Joko Lianto Buliali, M.Sc. yang telah membantu, membimbing dan member pengarahan kepada penulis selama mengikuti studi di MMT-ITS.
3. Yang Terhormat Para Dosen yang telah membimbing dan mengajar selama mengikuti studi di MMT-ITS.
4. Semua keluarga, teman-teman, dan pihak-pihak yang tidak dapat disebutkan satu persatu atas dukungan dan bantuannya kepada penulis.

Penulis telah berusaha sebaik-baiknya dalam mengerjakan Tesis ini, tetapi penulis menyadari bahwa tesis ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun dari pembaca demi pengembangan tesis ini ke arah yang lebih baik.

Surabaya, Januari 2015

Amelia Halim

IDENTIFIKASI MAHASISWA YANG MEMPUNYAI KECENDERUNGAN LULUS TIDAK TEPAT WAKTU PADA PROGRAM STUDI MMT-ITS DENGAN MENGGUNAKAN ALGORITMA C4.5

Nama Mahasiswa : Amelia Halim
NRP : 9112205304
Pembimbing : Prof. Dr. Ir. Joko Lianto Buliali, MSc

ABSTRAK

Setiap tahun Universitas menerima mahasiswa dari berbagai program, lokasi yang berbeda, latar belakang pendidikan yang beraneka ragam dan berbagai nilai dalam ujian masuk. Selain itu juga berasal dari universitas dengan akreditasi yang berbeda, masing-masing universitas memiliki kurikulum pelajaran yang berbeda dan tingkat kedalaman mata pelajaran yang diberikan juga berbeda. Proses menganalisis masa lalu kinerja mahasiswa akan memberikan perspektif yang lebih baik dari kemungkinan kinerja mahasiswa di masa depan. Hal ini bisa dicapai dengan memanfaatkan data yang sudah ada dan mengelolanya menggunakan konsep data mining.

Penelitian ini ditujukan untuk mengidentifikasi mahasiswa yang mempunyai kecenderungan lulus tidak tepat waktu. Penelitian ini menggunakan data pendidikan berupa sampel data alumni dari mahasiswa lulusan MMT – ITS dari angkatan 2009-2012. Metode yang digunakan adalah algoritma C4.5 yang merupakan salah teknik pengklasifikasian dalam data mining. Penelitian ini akan melakukan analisis untuk memperoleh informasi tentang kecenderungan mahasiswa lulus tepat waktu atau lulus tidak tepat waktu.

Tingkat akurasi penelitian ini akan diperoleh dari jumlah kasus yang menunjukkan indikasi yang sesuai. Penelitian ini akan diuji menggunakan recall dan precision. Penelitian ini akan diuji menggunakan akurasi, recall dan precision. Uji coba dilakukan dalam 3 macam skenario. Skenario I mempunyai rata-rata nilai precision, recall dan akurasi yaitu 46.27%, 31.1 % , dan 60.35%. Skenario II mempunyai rata-rata nilai precision, recall dan akurasi yaitu 47.26%, 32.05%, dan 61.25%. Skenario II mempunyai rata-rata nilai precision, recall dan akurasi yaitu 47.61%, 35.39%, dan 61.50%.

Kata Kunci: Algoritma C4.5, Data Mining , Klasifikasi, Penelitian Pendidikan.

IDENTIFICATION OF STUDENTS WHO HAVE GRADUATED NOT IN TIME ON MMT-ITS USING C4.5 CLASSIFICATION ALGORITHM

By : Amelia Halim
Student Identify Number : 9112205304
Supervisor : Prof. Dr. Ir. Joko Lianto Buliali, M.Sc

ABSTRAK

Every year, educational institutes admit students under various courses from different locations, educational background and with varying merit scores in entrance examinations. Moreover, schools and junior colleges may be affiliated to different boards, each board having different subjects in their curricula and also different level of depths in their subjects. Analyzing the past performance of admitted students would provide a better perspective of the probable academic performance of students in the future. This can very well be achieved using the concepts of data mining.

This study aimed to identify students who have a tendency not to pass on time. This study uses the data in the form of education alumni data sample of graduate students MMT - ITS of force from 2009 to 2012. The method used is a C4.5 algorithm which is one of classification techniques in data mining. This study will conduct an analysis to obtain information about the tendency of students to graduate on time or do not pass on time.

The level of accuracy of this research will be obtained from the number of cases that show the appropriate indications. . This study will be tested using the recall and precision. The trial was conducted in 3 kinds of scenarios. Scenario I have an average value of precision, recall and accuracy is 46.27%, 31.1%, and 60.35%. Scenario II has an average value of precision, recall and accuracy is 47.26%, 32.05%, and 61.25%. Scenario II has an average value of precision, recall and accuracy is 47.61%, 35.39%, and 61.50%.

Keyword: C4.5 Algorithm, Data Mining , Classification, Educational Research.

DAFTAR ISI

ABSTRAK	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	vi
DAFTAR TABEL	vii
DAFTAR SEGMENT PROGRAM	viii
 BAB 1 PENDAHULUAN	 1
1.1. Latar Belakang	1
1.2. Perumusan Masalah	2
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	3
1.6 Sistematisa Penulisan Tesis	4
 BAB 2 TINJAUAN PUSTAKA	 5
2.1. Profil Umum MMT ITS	5
2.2 Data Mining	7
2.3 Preprocessing Data	10
2.4 Klasifikasi (Classification)	11
2.5 Algoritma C4.5	12
2.5.1 Algoritma Konstruksi Pohon	13
2.5.2 Komputasi Gain Ratio pada Konstruksi Pohon C4.5	13
2.6 Pre Pruning	14
2.7 HTML dan CSS	15
2.8 PHP	16
2.9 My SQL	16
 BAB 3 METODOLOGI PENELITIAN	 17
3.1 Pengumpulan Informasi dan Data yang Terkait	18
3.2 Mengolah Data dan Menentukan Variabel yang Akan Digunakan	18

3.3 Melakukan Data Mining dengan Menggunakan Algoritma C45	22
3.4 Membuat Rancangan Flowchart Sistem	24
3.5 Perancangan Basis Data	25
3.6 Pengujian dan Evaluasi Sistem	26
3.7 Penarikan Kesimpulan	26
BAB 4 ARSITEKTUR PROGRAM	27
4.1 Gambaran Arsitektur Program Utama	27
4.1.1 Gambaran Detail Arsitektur	27
4.1.2 Input dan Output program	28
4.2 Analisa Kebutuhan Sistem	29
4.3 Desain Tabel dan Struktur Data	30
4.3.1 Desain Tabel	30
4.3.1.1 Tabel User	30
4.3.1.2 Tabel Form Atribut	31
4.3.1.3 Tabel Atribut	32
4.3.1.4 Tabel Data Keputusan	33
4.3.1.5 Tabel Data Keputusan Kinerja	36
4.3.1.6 Tabel Data Penentu Keputusan	39
4.3.1.7 Tabel Data Training	40
4.3.1.8 Tabel Iterasi C45	44
4.3.1.9 Tabel Mining C45	46
4.3.1.10 Tabel Pohon Keputusan C45	48
4.3.1.11 Tabel Rule C45	49
4.3.1.12 Tabel Rule Penentu Keputusan	50
4.3.1.12 Tabel Akreditasi	51
4.3.1.12 Tabel Program Studi	52
4.3.2 Desain Struktur Data	53
4.4 Alur Kerja	53
BAB 5 IMPLEMENTASI DAN UJI COBA SISTEM	55
5.1 Implementasi Sistem	55

5.1.1 Implementasi Program	55
5.1.2 Implementasi Desain Interface.....	61
5.2 Uji Coba Sistem	67
5.2.1 Uji Coba Skenario I.....	67
5.2.2 Uji Coba Skenario II	69
5.2.3 Uji Coba Skenario III.....	71
5.2.4 Perbandingan Hasil Uji Coba Skenario I , Skenario II, Skenario III	74
5.2.5 Uji Coba Program Studi MTI.....	75
5.2.5 Uji Coba Program Studi MI.....	76
5.2.5 Uji Coba Program Studi MP	77
BAB 6 KESIMPULAN DAN SARAN	79
6.1 Kesimpulan	79
6.2 Saran.....	80
DAFTAR PUSTAKA	81
LAMPIRAN A DATA INPUT	A-1
LAMPIRAN B CONTOH OUTPUT	B-1
LAMPIRAN C PERHITUNGAN GAIN.....	C-1

DAFTAR GAMBAR

Gambar 2.1 Proses Knowlegde Discovery in Database(KDD).....	9
Gambar 3.1 Tahapan Metodologi Penelitian.....	17
Gambar 3.2 Flowchart System	24
Gambar 3.3 Simbol – Simbol dalam DFD	25
Gambar 3.4 Perhitungan Recall, Precision dan Accuracy	26
Gambar 4.1 Arsitektur Program Utama	28
Gambar 4.2 Alur Kerja	54
Gambar 5.1 Halaman Login	60
Gambar 5.2 Halaman Website Utama	60
Gambar 5.3 Sub Menu Data Training	61
Gambar 5.4 Sub Menu Partisi Data.....	62
Gambar 5.5 Sub Menu Proses Mining C4.5.....	62
Gambar 5.6 Sub Menu Perhitungan C4.5.....	63
Gambar 5.7 Sub Menu Kinerja Perbandingan.....	63
Gambar 5.8 Sub Menu Kinerja Tabel Penilaian.....	64
Gambar 5.9 Sub Menu Penentu Keputusan.....	64
Gambar 5.10 Accuracy, Recall, dan Precision.....	74
Gambar 5.11 Perbandingan skenario I, skenario II, Skenario III.....	74
Gambar 5.12 Perbandingan Program Studi MI, MTI, MP dan ALL.....	78

DAFTAR TABEL

Tabel 3.1 Tabel List Jurusan yang Sebidang Per Program Studi S2.....	19
Tabel 3.2 Tabel Kel IPK, GMAT, MBID, Wawancara, TOEFL, SAKhir	21
Tabel 4.1 Struktur Table User.....	30
Tabel 4.2 Struktur Tabel Form Atribut.....	31
Tabel 4.3 Struktur Tabel Atribut.....	32
Tabel 4.4 Struktur Tabel Data Keputusan.....	33
Tabel 4.5 Struktur Tabel Data Keputusan Kinerja.....	36
Tabel 4.6 Struktur Tabel Data Penentu Keputusan.....	40
Tabel 4.7 Struktur Tabel Data Training.....	40
Tabel 4.8 Struktur Tabel Iterasi C45.....	44
Tabel 4.9 Struktur Tabel Mining C45.....	46
Tabel 4.10 Struktur Tabel Pohon Keputusan C45	48
Tabel 4.11 Struktur Tabel Rule C45	49
Tabel 4.12 Struktur Tabel Rule Penentu Keputusan.....	50
Tabel 4.12 Struktur Tabel Akreditasi.....	51
Tabel 4.12 Struktur Tabel Program Studi	52
Tabel 5.1 Hasil Precision Uji Coba Skenario I	68
Tabel 5.2 Hasil Recall Uji Coba Skenario I.....	68
Tabel 5.3 Hasil Accuracy Uji Coba Skenario I.....	69
Tabel 5.4 Hasil Precision Uji Coba Skenario II.....	70
Tabel 5.5 Hasil Recall Uji Coba Skenario II	70
Tabel 5.6 Hasil Accuracy Uji Coba Skenario II	71
Tabel 5.7 Hasil Precision Uji Coba Skenario III	72
Tabel 5.8 Hasil Recall Uji Coba Skenario III	72
Tabel 5.9 Hasil Accuracy Uji Coba Skenario III	73
Tabel 5.10 Hasil Uji Coba Program Studi MTI	75
Tabel 5.11 Hasil Uji Coba Program Studi MI	76
Tabel 5.12 Hasil Uji Coba Program Studi MP.....	77



DAFTAR LAMPIRAN

Lampiran A Data Input A1-A16

Lampiran B Contoh Data Output B1-B3

Lampiran C Contoh Perhitungan Gain Dan Entrophy C1- C5

DAFTAR SEGMENT PROGRAM

Segmen Program 5.1 PerhitunganC45 Part I.....	51
Segmen Program 5.2 Function perhitunganC45 Part II	53
Segmen Program 5.3 Function perhitunganC45 Part III	54
Segmen Program 5.4 Function perhitunganC45 Part IV	54
Segmen Program 5.5 Function perhitunganC45 Part V	55
Segmen Program 5.6 Function insertAtributPohonKeputusan.....	55
Segmen Program 5.7 Function getInfGainMax.....	56
Segmen Program 5.8 Function loopingPerhitungaPrePruning.....	57
Segmen Program 5.9 Function PerhitunganKinerja.....	57
Segmen Program 5.10 Function PenentuanKeputusan	59

BAB I

PENDAHULUAN

Pada bab pendahuluan akan dijelaskan mengenai latar belakang, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan tesis. Latar belakang membahas tentang latar belakang yang mendasari penelitian ini di lakukan.

1.1 Latar Belakang

Kualitas perguruan tinggi, khususnya program studi di Indonesia diukur berdasarkan akreditasi yang dilaksanakan oleh Badan Akreditasi Nasional Perguruan Tinggi atau BAN PT. Menurut BAN PT (BAN PT., 2011) kualitas tersebut diukur berdasarkan 7 standar utama, salah satu nya adalah Mahasiswa dan Lulusan. Khusus mengenai evaluasi standar mahasiswa dan lulusan, komponen yang dinilai adalah: sistem rekrutmen mahasiswa baru, dan lulusan (rata- rata masa studi dan IPK). Berdasarkan uraian diatas, maka dapat diambil kesimpulan bahwa kualitas sebuah perguruan tinggi salah satunya ditentukan oleh rekrutmen mahasiswa baru dan lama masa studi mahasiswa.

Salah satu permasalahan utama dari institusi perguruan tinggi adalah untuk meningkatkan kualitas pendidikan untuk mahasiswa dan untuk meningkatkan kualitas dari keputusan manajerial institusi. Salah satu cara untuk mencapai kualitas level mutu tertinggi dari sistem perguruan tinggi adalah dengan menggali pengetahuan dari data bidang pendidikan sebagai atribut pembelajaran utama yang mempengaruhi pencapaian mahasiswa (Abu., et al, 2012). Data-data bidang pendidikan pada umumnya bisa berupa data profile mahasiswa, mata kuliah, KRS (kartu rencana studi), data alumni, dan sebagainya, yang biasanya tersimpan dalam database Sistem Informasi Akademik Kampus (SIAK) dalam jumlah yang besar, dimana sebenarnya dari data bidang pendidikan tersebut dapat digunakan untuk menggali sebuah informasi.

Setiap tahun, lembaga pendidikan menerima mahasiswa dalam berbagai program dari lokasi yang berbeda, latar belakang pendidikan dan dengan berbagai nilai dalam ujian masuk. Selain itu juga berasal dari universitas dengan akreditasi

yang berbeda, masing-masing universitas memiliki berbeda pelajaran dalam kurikulum dan tingkat juga berbeda pada kedalaman mata pelajaran yang diberikan. Menganalisis masa lalu kinerja mahasiswa akan memberikan perspektif yang lebih baik dari kemungkinan kinerja mahasiswa di masa depan. Hal ini bisa sangat baik dicapai dengan menggunakan konsep data mining.

Penemuan pengetahuan dalam database (Knowledge Discovery in Databases/KDD), sering disebut *Data Mining* (Penambangan Data), mengacu pada penemuan informasi yang berguna dari kumpulan data yang besar (Goela, et al, 2012). Dengan memanfaatkan data mining pada data bidang pendidikan, sebuah institusi perguruan tinggi bisa memperoleh suatu informasi yang berguna, dimana selanjutnya informasi tersebut dapat menjadi suatu landasan untuk melakukan perbaikan untuk meningkatkan kualitas perguruan tinggi.

Penelitian ini menggunakan sampel data alumni dari mahasiswa lulusan MMT – ITS dari angkatan 2009-2012. Metode yang digunakan adalah algoritma C4.5. Algoritma C4.5 adalah algoritma klasifikasi data dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan. Kelebihan algoritma C4.5 dapat mengolah data numerik (kontinyu) dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan tercepat diantara algoritma-algoritma yang lain.

Penelitian ini ditujukan untuk membangun aplikasi yang memberikan *warning* kepada dosen wali terhadap kecenderungan seorang mahasiswa lulus tidak tepat waktu pada awal masa studi mahasiswa. Diharapkan dari penelitian yang dilakukan terhadap sampel data alumni tersebut dapat diperoleh suatu informasi yang bisa membantu pihak institusi pendidikan untuk merancang strategi meningkatkan kualitas perguruan tingginya.

1.2 Perumusan Masalah

Sesuai dengan latar belakang yang ada maka permasalahan pada universitas MMT – ITS adalah membuat suatu aplikasi prediksi lama masa studi mahasiswa yang mampu menangani masalah :

1. Faktor-faktor apa yang mempengaruhi kecenderungan mahasiswa lulus tidak tepat waktu?

2. Bagaimana membuat suatu aplikasi yang yang mampu mendeteksi kecenderungan mahasiswa lulus tidak tepat waktu?

1.3 Batasan Masalah

Aplikasi yang dikembangkan memberikan peringatan dini tentang kecenderungan mahasiswa lulus tidak tepat waktu pada awal masa kuliah dan tidak menangani pemberian solusi manajerial untuk mengatasi permasalahan mahasiswa tersebut.

1.4 Tujuan Penelitian

Sesuai dengan perumusan masalah yang ada maka tujuan penelitian adalah membangun sebuah aplikasi yang mampu menangani masalah:

1. Menganalisa faktor – faktor yang mempengaruhi kecenderungan mahasiswa lulus tidak tepat waktu.
2. Membuat aplikasi yang mampu mendeteksi kecenderungan mahasiswa lulus tidak tepat waktu.

1.5 Manfaat Penelitian

Sesuai dengan rumusan masalah dan tujuan yang telah disebutkan, maka penelitian ini diharapkan berguna bagi Universitas, bagi pengembang ilmu pengetahuan, dan bagi penulis sendiri.

1. Bagi Universitas

Penelitian ini diharapkan membantu menyelesaikan permasalahan universitas dalam mendeteksi dini dan menangani mahasiswa yang memiliki kecenderungan lulus tidak tepat waktu sehingga universitas bisa mempertahankan akreditasinya.

2. Bagi Ilmu Pengetahuan

Bagi Ilmu Pengetahuan, hasil dari penelitian ini diharapkan dapat digunakan untuk menyelesaikan permasalahan yang sama atau mirip dengan studi kasus yang ada dan dikembangkan sesuai dengan disiplin ilmu yang dimiliki baik oleh para akademisi maupun praktisi

1.6 Sistematika Penulisan Tesis

Sistematika yang digunakan dalam penulisan tesis ini adalah sebagai berikut:

BAB I PENDAHULUAN

Berisi Latar Belakang yang mendorong dilakukannya penelitian, perumusan masalah serta tujuan dan manfaat penelitian.

BAB II TINJAUAN PUSTAKA

Membahas tentang kondisi lokus organisasi, dasar teori, temuan, dan bahan penelitian sebelumnya yang diperoleh dari berbagai referensi yang dijadikan dasar melakukan penelitian.

BAB III METODOLOGI PENELITIAN

Pada bagian ini diuraikan desain, metoda, atau pendekatan yang akan digunakan dalam menjawab permasalahan penelitian / studi untuk mencapai studi penelitian, serta tahapan penelitian secara rinci, singkat, dan jelas.

BAB IV ARSITEKTUR PROGRAM

Pada bagian ini akan diuraikan tentang arsitektur program utama. Penjelasan arsitektur program meliputi tentang gambaran detail arsitektur, input dan output program, analisa kebutuhan sistem, desain tabel dan struktur data, serta alur kerja dari program..

BAB V IMPLEMENTASI, UJICOBA, DAN EVALUASI SISTEM

Berisi hasil implementasi sistem ke dalam bentuk perangkat lunak yang dapat memberikan peringatan dini (*early warning system*), ujicoba sistem, evaluasi kerja sistem.

BAB VI KESIMPULAN DAN SARAN

Berisi kesimpulan dari pembuatan program pada Tesis ini. Pada bab ini juga disertakan saran mengenai kemungkinan pengembangan pada program yang dibuat.

BAB II

TINJAUAN PUSTAKA

Pada bab ini akan dibahas tentang tinjauan pustaka akan dipakai dalam pembuatan aplikasi pendeteksian mahasiswa yang mempunyai kecenderungan lulus tidak tepat waktu. Adapun dasar Teori yang akan dibahas adalah KDD(*Knowledge Discovery in Database*), *Preprocessing*, Klasifikasi(*Classification*), dan algoritma C4.5.

2.1 Profil Umum MMT ITS

Institut Teknologi Sepuluh Nopember (ITS) telah memberi kontribusi yang signifikan dalam kemajuan teknologi dan telah ikut melahirkan pemimpin, pengusaha dan manajer yang handal secara nasional. Tetapi tantangan globalisasi, tekanan hiper kompetisi di dunia yang semakin kompleks dan bergantung itu tidak cukup solusinya hanya dengan mengandalkan kepada kemajuan di bidang teknologi. Sebagai teknopreneurship yang handal, tantangan ini memerlukan respons yang cepat, unik, dinamis dan kompleks dalam dunia bisnis. Untuk itu para manajer dan pemimpin harus memiliki pengetahuan bisnis dan entrepreneurship yang baik, mempunyai kompetensi dalam mencari solusi manajerial terbaik secara kreatif, efektif, dan realistis untuk tiap situasi.

Untuk menghadapi tantangan di atas, diperlukan pengetahuan lintas dan multi disiplin. Oleh karena itu, semenjak tahun 1996 ITS telah menyelenggarakan Program Studi Magister Manajemen Teknologi (MMT) yang berada di bawah pengelolaan Program Pascasarjana ITS. Program pendidikan yang diselenggarakan oleh MMT-ITS berorientasi praktis, komprehensif, dan relevan dalam menyiapkan calon-calon manajer, pelaku bisnis dan industri, pemimpin dan peneliti bisnis dan industri.

MMT-ITS meliputi bidang keilmuan yang lintas dan multi disiplin, dan program pendidikan ini dirancang untuk membekali para manajer, pemimpin dan eksekutif profesional dengan pengetahuan dan kompetensi serta ketrampilan manajerial. Kompetensi tersebut diperlukan dalam menghadapi perubahan serta

dinamika perkembangan sistem organisasi dan bisnis, perusahaan, industri yang sangat dipengaruhi oleh perubahan dan kemajuan teknologi, ekonomi dan keuangan.

Sebagai Perguruan Tinggi Nasional terkemuka di Indonesia, ITS memiliki banyak tenaga akademis dengan latar pendidikan pascasarjana dari perguruan tinggi luar negeri terbaik, serta mempunyai pengalaman praktis di dunia industri dan bisnis. Sumber daya insani yang dimiliki oleh ITS ini memberikan jaminan kualitas pendidikan serta proses belajar mengajar di MMT-ITS.

MMT-ITS memiliki visi sebagai pusat pendidikan berjenjang S2 unggulan di Bidang Manajemen Teknologi dengan mengembangkan sumber daya insani agar berkemampuan mengintegrasikan strategi teknologi dan manajerial untuk meningkat daya saing organisasi serta memiliki sikap dan perilaku profesionalism.

Sasaran kompetensi yang diharapkan dari proses belajar mengajar di MMT-ITS antara lain:

1. Kemampuan mengembangkan dan memutakhirkan wawasan bisnis, ketrampilan manajerial dan merumuskan strategi teknologi
2. Kemampuan manajerial untuk mengelola organisasi, industri, bisnis dan perusahaan secara mandiri, efektif dan efisien
3. Kemampuan memecahkan masalah organisasi dan industri yang kompleks berbasis prinsip-prinsip manajemen dan bisnis dengan etis.

Berbagai program kerjasama pendidikan telah dilaksanakan oleh Program Studi MMT-ITS dengan berbagai Instansi Pemerintah dan Industri. Pada tahun 1996 sampai 2001 Program Studi MMT-ITS mengadakan kerjasama dengan Pemerintah Daerah Tingkat I Papua dalam menyelenggarakan pendidikan Bidang Manajemen Proyek, Manajemen Pembangunan Kota, Manajemen Rekayasa Lingkungan dan Manajemen Teknologi Transportasi. Pada tahun 2003, MMT-ITS melaksanakan Program kerjasama pendidikan dengan PT Chevron Pacific Indonesia (C/O. PT CALTEX Pacific Indonesia) dengan konsentrasi Bidang Keahlian Manajemen Teknologi Informasi. Beberapa kerjasama yang telah dilaksanakan antara lain dengan BPKP (Bidang Keahlian Manajemen Teknologi Informasi), dengan Departemen Perhubungan (Bidang Keahlian Manajemen

Teknologi Transportasi dan Manajemen Teknologi Informasi), dan dengan PT Paiton (Bidang Keahlian Manajemen Industri).

2.2 Data Mining

Data Mining memang salah satu cabang ilmu komputer yang relatif baru dan sampai sekarang orang masih memperdebatkan untuk menempatkan data mining di bidang ilmu mana, karena data mining menyangkut database, kecerdasan buatan (artificial intelligence), statistik, dsb. Ada pihak yang berpendapat bahwa data mining tidak lebih dari machine learning atau analisa statistik yang berjalan di atas database. Namun pihak lain berpendapat bahwa database berperan penting di data mining karena data mining mengakses data yang ukurannya besar (bisa sampai terabyte) dan disini terlihat peran penting database terutama dalam optimisasi query-nya.

Banyak istilah yang digunakan untuk menunjukkan proses data mining (contoh : knowledge discovery, knowledge extraction, data/pattern analysis, data archeology, data dredging, information harvesting, business intelligence, dll).

Untuk dapat memberikan pemahaman tentang data mining, penulis akan mencoba menguraikan urutan fakta yang terjadi, agar dapat memberikan persepsi tersendiri. Berikut ini adalah beberapa fakta yang terjadi :

- Banyak sekali organisasi, baik dari dunia bisnis ataupun pemerintah berurusan dengan sejumlah sumber informasi dan juga pengelolaan basis data informasi tersebut, dan bukan tidak mungkin termasuk di dalamnya kebutuhan akan pembangunan data warehouse dalam skala besar.
- Data yang tersimpan tidak dapat secara langsung di analisa dengan metode metode statistik standar. Hal ini disebabkan karena adanya beberapa rekord yang hilang ataupun juga karena data nya dalam dimensi ukuran kualitatif dan bukan kuantitatif.
- Karena tingkat pertumbuhan ukuran basis data yang sangat cepat, bahkan terkadang sistem administrator nya sendiri pun mengalami kendala untuk mengetahui informasi yang terkandung di dalamnya atau sekedar mengetahui hubungannya dengan pertanyaan pertanyaan yang timbul.

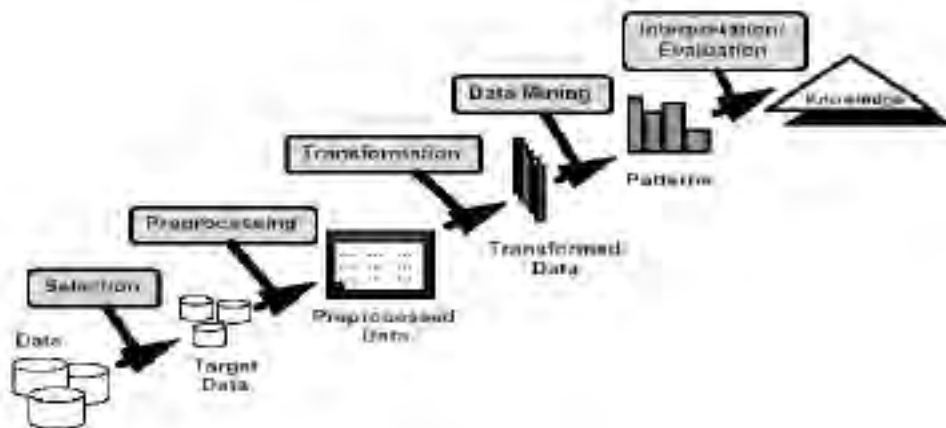
- Akan menjadi suatu keuntungan tersendiri apabila suatu organisasi mempunyai cara untuk “menggali” sumber informasi nya yang berupa basis data yang besar, sehingga dapat diketahui informasi yang penting dan juga pola pola yang kemungkinan terkandung di dalamnya.
- Sekarang terdapat beberapa metodologi data mining yang kemungkinan dapat berguna untuk menganalisa sumber sumber data dalam rangka menemukan pola dan tren terbaru.

Definisi sederhana dari data mining adalah ekstraksi informasi atau pola yang penting atau menarik dari data yang ada di database yang besar. Dalam jurnal ilmiah, data mining juga dikenal dengan nama Knowledge Discovery in Databases (KDD).

Sesuai yang tercantum dalam buku “Advances in Knowledge Discovery dan Data mining” terdapat definisi sebagai berikut:

Knowledge discovery (data mining) in databases (KDD) adalah keseluruhan proses non-trivial untuk mencari dan mengidentifikasi pola (pattern) dalam data, dimana pola yang ditemukan bersifat sah (valid), baru (novel), dapat bermanfaat (potentially usefull), dapat dimengerti (ultimately understandable)[2].

Istilah data mining dan knowledge discovery in databases (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda akan tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining. KDD (Knowledge Discovery in Database) merupakan keseluruhan proses konversi data mentah menjadi pengetahuan yang bermanfaat yang terdiri dari serangkaian tahap transformasi meliputi data preprocessing dan postprocessing.



Gambar 2.1
Proses Knowledge Discovery in Database(KDD)

Berdasarkan gambar di atas proses KDD secara garis besar dapat dijelaskan sebagai berikut:

a. Data Selection

Pemilihan (selection) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

b. Pre-processing

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus KDD. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

c. Transformation

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

d. Data mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

e. Interpretation

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut dengan interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

Proses KDD secara garis besar memang terdiri dari 5 tahap seperti yang telah dijelaskan sebelumnya. Akan tetapi, dalam proses KDD yang sesungguhnya, dapat saja terjadi iterasi atau pengulangan pada tahap tahap tertentu. Pada setiap tahap dalam proses KDD, seorang analis dapat saja kembali ke tahap sebelumnya. Sebagai contoh, pada saat coding atau data mining, analis menyadari proses cleaning belum dilakukan dengan sempurna, atau mungkin saja analis menemukan data atau informasi baru untuk “memperkaya” data yang sudah ada.

2.3 Preprocessing Data

Data Preparation atau bisa disebut juga dengan data preprocessing adalah suatu proses/langkah yang dilakukan untuk membuat data mentah menjadi data yang berkualitas(input yang baik untuk data mining tools).

Data harus di preprocessing terlebih dahulu disebabkan dalam data mentah masih terdapat data yang :

- incomplete, yaitu data yang kekurangan nilai atribut atau hanya mengandung agregat data (contoh : address = " ").
- noisy, yaitu data yang masih mengandung error dan outliers (contoh : salary = -10).
- inconsistent, yaitu data yang mengandung discrepansi dalam code dan nama atau singkatnya datanya tidak konsisten (contoh : dulu rating = 1,2,3 sekarang a,b,c):.

Langkah-Langkah dalam data preparation sebagai berikut :

1. Data Cleaning

Dalam data cleaning yang akan kita lakukan antara lain mengisi missing value, mengidentifikasi outlier, menangani data noise, mengoreksi data yang tidak konsisten, dan menyelesaikan masalah redudansi data akibat integrasi data.

2. Data Integration

Data integration adalah suatu langkah untuk menggabungkan data dari beberapa sumber. Data integration hanya dilakukan jika data berasal dari tempat yang berbeda-beda (sumber data tidak hanya dari 1 tempat). Langkah yang dilakukan antara lain mengintegrasikan skema, mengidentifikasi masalah entitas, dan mendeteksi sekaligus menyelesaikan konflik pada nilai data.

3. Data Transformation

Data transformation yaitu mengubah suatu data supaya diperoleh data yang lebih berkualitas. Yang akan dilakukan antara lain menghilangkan noise dari data (smoothing), meng-agregasi data, generalisasi data, normalisasi data, dan pembentukan atribut/fitur.

4. Data Reduction

Data Reduction yaitu langkah untuk mereduksi dimensi, atribut ataupun jumlah data. Yang akan dilakukan antara lain agregasi data cube, reduksi dimensi, diskretisasi, dan kompresi data.

2.4 Klasifikasi (*Classification*)

Classification adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri bisa berupa aturan “jika-maka”, berupa decision tree, formula matematis atau neural network.

Decision tree adalah salah satu metode classification yang paling populer karena mudah untuk diinterpretasi oleh manusia. Disini setiap percabangan menyatakan kondisi yang harus dipenuhi dan tiap ujung pohon menyatakan kelas data. Algoritma decision tree yang paling terkenal adalah C4.5, tetapi akhir-akhir ini telah dikembangkan algoritma yang mampu menangani data skala besar yang tidak dapat ditampung di main memory seperti RainForest. Metode-metode

classification yang lain adalah Bayesian, neural network, genetic algorithm, fuzzy, case-based reasoning, dan k-nearest neighbor.

Proses classification biasanya dibagi menjadi dua fase : learning dan test. Pada fase learning, sebagian data yang telah diketahui kelas datanya diumpangkan untuk membentuk model perkiraan. Kemudian pada fase test model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tsb. Bila akurasinya mencukupi model ini dapat dipakai untuk prediksi kelas data yang belum diketahui.

2.5 Algoritma C4.5

Algoritma C4.5 mengkonstruksi pohon keputusan dari data pelatihan, yang berupa kasus-kasus atau rekord-rekord (tupel) dalam basisdata. Setiap kasus berisikan nilai dari atribut-atribut untuk sebuah kelas. Setiap atribut dapat berisi data diskret atau kontinyu numerik). C4.5 juga menangani kasus yang tidak memiliki nilai untuk sebuah atau lebih atribut. Akan tetapi, atribut kelas hanya bertipe diskret dan tidak boleh kosong .

Tiga prinsip kerja algoritma C4.5 adalah:

- Pertama, konstruksi pohon keputusan. Tujuan dari algoritma konstruksi pohon keputusan adalah mengkonstruksi struktur data pohon (dinamakan pohon keputusan) yang dapat digunakan memprediksi kelas dari sebuah kasus atau rekord.
- Kedua, pemangkasan pohon ke putusan dan evaluasi. Karena pohon yang dikonstruksi dapat berukuran besar dan tidak mudah “dibaca”, C4.5 dapat menyederhanakan pohon dengan melakukan pemangkasan berdasarkan nilai tingkat kepercayaan. Pemangkasan juga bertujuan untuk mengurangi tingkat kesalahan prediksi pada kasus (rekord) baru.
- Ketiga, pembuatan aturan-aturan dari pohon keputusan. Aturan-aturan dalam bentuk if-then diturunkan dari pohon ke putusan dengan melakukan penelusuran dari akar sampai ke daun.

2.5.1 Algoritma Konstruksi Pohon

Algoritma dasar untuk induksi pohon keputusan pada C4.5 adalah algoritma greedy yang membangun pohon keputusan dari atas ke bawah (top-down) secara rekursif dengan cara divide dan conquer. Masukan dari algoritma ini adalah set data yang berisi sampel-sampel data dan kandidat atribut yang harus ditelaah, terdiri dari minimal sebuah atribut prediktor dan sebuah atribut kelas. Atribut prediktor dapat bertipe diskret atau numerik, sedangkan atribut kelas harus bertipe diskret. Dalam terminologi basisdata, set data ini berupa tabel, sedangkan sampel adalah record. Set data ini dapat memiliki atribut (kolom tabel) bertipe diskret maupun kontinyu. Adapun langkah-langkah konstruksi pohon ditunjukkan pada Algoritma 2.1.

Algoritma 2.1: Algoritma konstruksi pohon keputusan

Narasi: Membuat pohon keputusan dari data pelatihan yang diberikan.

Masukan: Sampel data pelatihan, samples, yang direpresentasikan dengan atribut bernilai diskret, kandidat himpunan atribut, attribute-list.

Keluaran: Sebuah pohon keputusan.

Metoda:

- (1) buat sebuah simpul N,
- (2) if samples memiliki kelas yang sama, C, then
- (3) return N sebagai simpul daun dengan label kelas C;
- (4) if attribute-list kosong then
- (5) return N sebagai simpul daun dengan label kelas terbanyak di samples
- (6) pilih test-attribute, yaitu salah satu atribut dari attribute-list dengan gain ratio terbesar;
- (7) beri label pada simpul N dengan test-attribute;
- (8) for setiap nilai a_i pada test-attribute;
- (9) tambahkan cabang pada simpul N untuk kondisi test-attribute = a_i ;
- (10) buat partisi sampel s_i dari samples dimana test-attribute = a_i ;
- (11) if s_i kosong then
- (12) tempelkan daun yang diberi label dengan kelas terbanyak di samples;

(13) else tempelkan simpul yang dibuat oleh Generate_decision_tree (si, attribute-list- test-attribute);

2.5.2 Komputasi Gain Ratio pada Kontruksi Pohon C4.5

Pada konstruksi pohon C4.5, di setiap simpul pohon, atribut dengan nilai gain ratio yang tertinggi dipilih sebagai atribut test atau split untuk simpul. Rumus dari gain ratio adalah $\text{gain ratio}(a) = \text{gain}(a) / \text{split info}(a)$, dimana $\text{gain}(a)$ adalah information gain dari atribut a untuk himpunan sampel X dan $\text{split info}(a)$ menyatakan informasi potensial yang didapat pada pembagian X menjadi n sub himpunan berdasarkan telaahan pada atribut a . Sedangkan $\text{gain}(a)$ didefinisikan sebagai [9]

$$\text{gain}(a) = \text{info}(X) - \text{info}_a(X) \quad (1)$$

dimana $\text{info}(X) = -\sum_{j=1}^k \frac{\text{freq}(C_j, X)}{|X|} \times \log_2 \left(\frac{\text{freq}(C_j, X)}{|X|} \right)$ dengan k adalah jumlah kelas pada himpunan rekord X . $\text{freq}(C_j, X)$ menyatakan jumlah sampel pada X yang memiliki nilai kelas C_j . $|X|$ menyatakan kardinalitas (jumlah anggota) himpunan

data X . $\text{info}_a(X) = \sum_{i=1}^n \frac{|X_i|}{|X|} \times \text{info}(X_i)$ menyatakan $\text{info}(X)$ dengan a adalah atribut yang ditelaah dan n

adalah jumlah sub himpunan yang dibentuk dari X (pada atribut diskret, n adalah jumlah nilai disting pada a , sedangkan pada atribut kontinyu, $n = 2$). Sedangkan

rumus $\text{split info}(a)$ adalah [9]: $\text{split info}(a) = -\sum_{i=1}^n \frac{|X_i|}{|X|} \times \log_2 \left(\frac{|X_i|}{|X|} \right) \quad (2)$ dimana

X_i menyatakan sub himpunan ke- i pada sampel X . Bahasan detil dari komputasi gain ratio ini dapat ditemukan di [4,9].

2.6 Pre Pruning

Pre pruning yaitu pemangkasan yang dilakukan sejak awal pembentukan pohon dengan cara menghentikan pembangunan suatu subtree lebih awal, yaitu dengan memutuskan untuk tidak lebih jauh mempartisi data training. Cara kerja pre pruning adalah dengan menghitung dulu nilai information gain untuk mengetahui nilai parent dan child. Setelah parent dan child diketahui kemudian dihitung nilai errornya, jika nilai error child lebih kecil parent maka parent

membentuk subtree lagi, tapi sebaliknya jika nilai error child lebih besar dari parent maka pruning dilakukan dan pembentukan subtree berhenti. Untuk menghitung nilai error digunakan rumus dibawah ini.

Rumus pre pruning :

$$e = \frac{r + \frac{z^2}{2n} + z \sqrt{\frac{r}{n} - \frac{r^2}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}$$

Dimana:

r = nilai perbandingan error rate

n = total sample

$z = \Phi^{-1}(c)$

c = confidence level

2.7 HTML dan CSS

HyperText Markup Language (HTML) adalah sebuah bahasa markup untuk membuat halaman web atau menampilkan informasi di web browser. HTML memungkinkan gambar dan benda untuk dimasukkan dan dapat digunakan untuk membuat bentuk-bentuk interaktif. Dari sini, dokumen terstruktur yang dibuat dengan menggunakan semantik struktural untuk teks seperti judul, link, daftar, paragraf, kutipan dll

CSS (Cascading Style Sheets) dirancang untuk memungkinkan pemisahan antara konten dokumen (dalam HTML atau bahasa markup yang sama) dan presentasi dokumen. Teknik ini digunakan untuk meningkatkan aksesibilitas isi juga untuk memberikan lebih banyak fleksibilitas dan kontrol dalam spesifikasi konten dan presentasi karakteristik. Hal ini memungkinkan beberapa halaman untuk berbagi format dan mengurangi redudansi.

2.8 PHP

PHP adalah singkatan dari "PHP: Hypertext Preprocessor", yaitu bahasa pemrograman yang digunakan secara luas untuk penanganan pembuatan dan pengembangan sebuah situs web dan bisa digunakan bersamaan dengan HTML..

PHP adalah bahasa pemrograman script server-side yang didesain untuk pengembangan web, tetapi juga bisa digunakan sebagai bahasa pemrograman umum (wikipedia). PHP pertama kali dikembangkan pada tahun 1995 oleh Rasmus Lerdorf, namun sekarang dikelola oleh The PHP Group. Situs resmi PHP beralamat di <http://www.php.net>.

Pada awalnya PHP adalah singkatan dari Personal Home Page, namun karena dalam perkembangannya PHP tidak hanya digunakan untuk membuat halaman web pribadi, PHP saat ini merupakan singkatan dari PHP: Hypertext Preprocessor, sebuah kepanjangan rekursif, yakni permainan kata dimana kepanjangannya berisi juga singkatan itu sendiri.

PHP dirilis dalam lisensi PHP License, sedikit berbeda dengan lisensi GNU General Public License (GPL) yang biasa digunakan untuk proyek Open Source. Namun penggunaan PHP tetap tidak dikenakan biaya (gratis).Kemudahan dan kepopuleran PHP sudah menjadi standar bagi programmer web di seluruh dunia. Menurut wikipedia pada februari 2014, sekitar 82% dari web server di dunia menggunakan PHP, dan menjadi dasar dari program CMS (Content Management System) populer seperti Joomla, Drupal, dan WordPress.

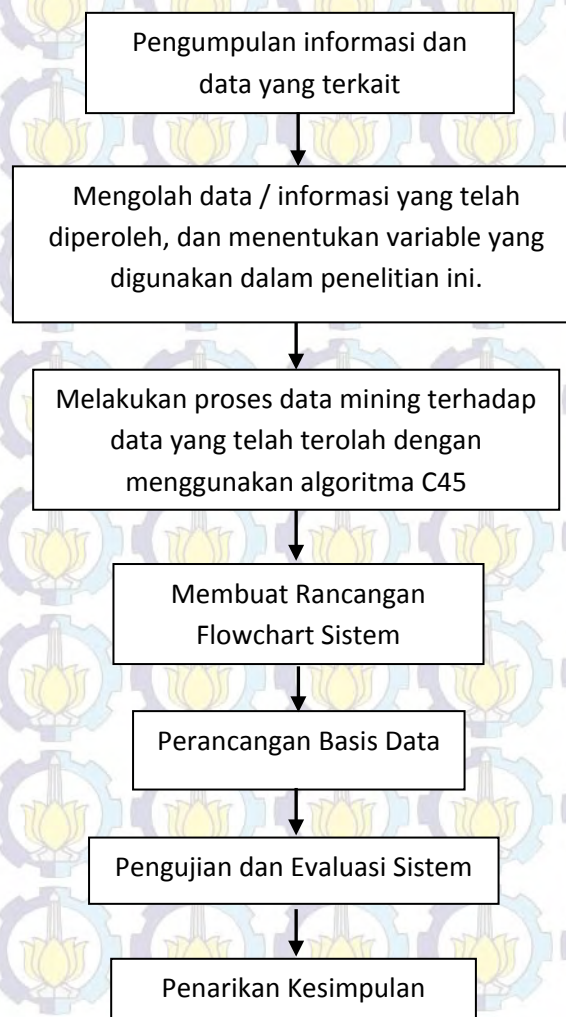
2.9 MySQL

MySQL adalah sebuah perangkat lunak sistem manajemen basis data SQL (database management system) atau DBMS yang multithread, multi-user,dengan sekitar 6 juta instalasi di seluruh dunia.

MySQL sebenarnya merupakan turunan salah satu konsep utama dalam database sejak lama, yaitu SQL (Structured Query Language). SQL adalah sebuah konsep pengoperasian database, terutama untuk pemilihan atau seleksi dan pemasukan data, yang memungkinkan pengoperasian data dikerjakan dengan mudah secara otomatis.

BAB III METODOLOGI PENELITIAN

Berikut adalah proses penelitian yang digunakan untuk pengidentifikasi mahasiswa yang lulus tidak tepat waktu. Penelitian akan dilaksanakan sesuai dengan prosedur dan langkah seperti pada diagram alur kerangka kerja pada gambar 3.1 sebagai berikut:



Gambar 3.1
Tahapan Metodologi Penelitian

3.1 Pengumpulan Informasi dan Data yang Terkait

Langkah awal adalah mencari informasi dari pihak MMT ITS bagian akademik mengenai data histori mahasiswa. Informasi tersebut didapatkan dengan mengumpulkan data – data mahasiswa yang pernah mengenyam pendidikan di MMT ITS. Data – data tersebut dikumpulkan menjadi satu untuk dijadikan bahan penelitian.

3.2 Mengolah Data dan Menentukan Variabel yang Akan Digunakan

Setelah data terkumpul, langkah selanjutnya adalah mengolah data tersebut menjadi bentuk data yang dibutuhkan, membersihkan data, dan menyeleksi data yang akan digunakan sebagai inputan proses data mining. Selain itu, dibutuhkan juga untuk menentukan variabel yang akan digunakan sebagai inputan proses data mining. Variabel – variabel tersebut yang akan digunakan untuk membangun pohon keputusan yang dibentuk oleh algoritma C45.

Pada Penelitian ini digunakan data sebanyak 178 data mahasiswa lulusan MMT – ITS dari angkatan 2009-2012. Dari proses pengumpulan data diperoleh 19 variabel meliputi NRP, Nama Perguruan Tinggi S1, Jurusan S1, IPK S1, Tahun Lulus S1, Tahun Masuk S2, Nama Perusahaan Tempat Bekerja, Pekerjaan, Nilai GMAT, Nilai Toefl, Nilai Materi Bidang, Nilai Wawancara, Nilai Skor Akhir, Program Studi S2, Bidang Keahlian S2, IPS Semester 1, IPS Semester 2, Sumber dana S2, dan Lama Studi. Variabel NRP digunakan sebagai Identifier karena sifatnya yang unik. Variable lama studi dalam satuan semester. Variable lama studi digunakan untuk menghasilkan class/Label dengan ketentuan lama studi ≤ 4 semester maka tergolong class Tepat Waktu, dan lama studi > 4 semester maka tergolong class Tidak Tepat Waktu.

Variabel Nama Perguruan Tinggi S1 mencatat nama Perguruan Tinggi S1. Sedangkan Jurusan S1 mencatat jurusan yang diambil mahasiswa saat S1. Variabel Nama Perguruan Tinggi S1 dan Jurusan S1 variabel mempunyai tingkat keanekaragaman yang tinggi. Pada tahap proses preprocessing data, Variabel Nama Perguruan Tinggi S1 dan variabel Jurusan S1 diproses membentuk satu variabel baru yaitu Akreditasi Perguruan Tinggi S1. Akreditasi Perguruan tinggi diperoleh dari website : <http://ban-pt.kemdiknas.go.id/direktori.php>.

Variabel Jurusan S1 dan Variabel Program Studi dipadukan untuk membentuk variabel baru yaitu Status Program Studi S1. Variabel Status Program Studi S1 mempunyai 2 macam nilai yaitu sebidang dan tidak sebidang. Variabel Status Program Studi S1 akan bernilai sebidang bila jurusan S1 dan program studi S2 merupakan satu bidang ilmu dan sebaliknya. Pada Table 3.1 dicantumkan list Jurusan S1 yang sebidang dengan variabel program studi S2.

Tabel 3.1
Tabel List Jurusan yang Sebidang Per Program Studi S2

Program Studi : MTI	Program Studi : MI
Sebidang:	Sebidang:
Teknik Informatika	Manajemen
Teknologi Informasi	Teknik Elektro
Sistem Informasi	Teknik Industri
Teknik Elektro	Teknik & Manajemen Industri
Manajemen Informatika & Komputer	Teknologi Industri Pertanian
	Teknik Mesin
	Ekonomi Manajemen
	Desain Produk Industri
Program Studi : MP	Marketing Manajemen
Sebidang:	Program Studi : MTK
Teknik Industri	Sebidang :
Teknik Sipil	Teknik Lingkungan
Arsitektur	Ilmu Lingkungan
Bisnis dan Manajemen	Program Studi : MBM
Teknik Perkapalan	Sebidang :
	Nautika
	Sistem Perkapalan
	Teknik Sipil

Variabel Tahun Lulus S1 adalah variabel yang berisikan Tahun seorang mahasiswa menyelesaikan kuliah S1. Dari 178 data sampel data terdapat 35 data yang variabel Tahun Lulus S1 kosong. Pada kasus ini peneliti, menggunakan nilai rata-rata Tahun Lulus S1 per angkatan untuk mengisi Variable Tahun Lulus S1 pada angkatan yang berkaitan. Variabel Tahun Masuk S2 dikurangi Variabel Tahun Lulus S1 diperoleh selisih tahun dari lulus S1 sampai melanjutkan studi S2. Setelah melakukan proses grouping nilai, dibentuklah variabel jarak S1-S2. Jarak

S1-S2 berisi 2 macam value yaitu dekat dan jauh. Jarak S1-S2 akan bernilai dekat apabila selisih tahun masuk S2 dan tahun lulus S1 kurang dari sama dengan 3 tahun, sedangkan Jarak S1-S2 akan bernilai jauh apabila selisih tahun masuk S2 dan tahun lulus S1 lebih dari 3 tahun.

Variabel Nama Perusahaan Tempat Bekerja adalah variabel yang menampung Nama Perusahaan tempat mahasiswa bekerja(jika telah bekerja). Sedangkan Variabel Pekerjaan adalah variabel yang menampung Jabatan mahasiswa tersebut di tempat bekerja. Variabel Nama Perusahaan Tempat Bekerja, dan Pekerjaan pada penelitian ini diabaikan karena missing value hampir 75% dari 178 data training yang ada.

Variabel IPK S1, IPK S2, Nilai GMAT, Nilai Toefl, Nilai Materi Bidang, Nilai Wawancara, Nilai Skor Akhir merupakan variabel yang tipenya numerik / angka. Pada preprocessing data, nilai dari tujuh variabel tersebut dikelompokkan menjadi 5 kelompok nilai / Kelas yaitu A, B, C, D, E. Pengelompokan dilakukan dengan cara mengenali terlebih dahulu sebaran data dari masing – masing variabel. Pertama – tama menentukan nilai minimum dan maximum dari sebuah variabel, lalu mencari panjang kelasnya dengan cara : $(\text{Nilai Maximum} - \text{Nilai Minimum}) / \text{jumlah kelas}$. Setelah itu menentukan batasan masing – masing kelas nilai. Tabel 3.1 menunjukan range dan kelas dari 7 variable nilai IPK S1, IPK S2, Nilai GMAT, Nilai Toefl, Nilai Materi Bidang, Nilai Wawancara, Nilai Skor Akhir.

Perhitungan Range IPK S1 menggunakan nilai min yaitu 2.5, nilai max yaitu 4, panjang kelas yaitu 0.3. Perhitungan Range IPS Semester 1 menggunakan nilai min yaitu 2.4, nilai max yaitu 4, panjang kelas yaitu 0.32. Perhitungan Range IPS Semester 2 menggunakan nilai min yaitu 2, nilai max yaitu 4, panjang kelas yaitu 0.4. Perhitungan Range [SAkhir,Wawancara] menggunakan nilai min yaitu 60, nilai max yaitu 100, panjang kelas yaitu 8. Perhitungan Range [GMAT , MBid] menggunakan nilai min yaitu 30, nilai max yaitu 100, panjang kelas yaitu 14. Nilai min [IPK S1, IPS Semester 1, IPS Semester 2,Sakhir, GMAT, MBID, Wawancara] adalah nilai terendah dari data sample yang ada, sedangkan nilai max [IPK S1, IPS Semester 1, IPS Semester 2, Sakhir, GMAT, MBID, Wawancara] adalah nilai tertinggi dari data sample yang ada. Perhitungan Range TOEFL

menggunakan nilai min yaitu 310, nilai max yaitu 677, panjang kelas yaitu 73.4. Nilai Min dan Nilai Max TOEFL menggunakan Nilai Max dan Nilai Min TOEFL versi PBT (*paper-based test*).

Tabel 3.2
Tabel Pengelompokan IPK, GMAT, MBID, Wawancara, TOEFL, SAKhir

IPK S1		SAKhir		GMAT	
Range	Kelas	Range	Kelas	Range	Kelas
4.00 - 3.79	A	100 - 92	A	100 - 86	A
3.78 - 3.58	B	91 - 84	B	85 - 72	B
3.57 - 3.37	C	83 - 76	C	71 - 58	C
3.36 - 3.16	D	75 - 68	D	57 - 44	D
< 3.16	E	< 68	E	< 44	E

MBID		Wawancara		TOEFL	
Range	Kelas	Range	Kelas	Range	Kelas
100 - 86	A	100 - 92	A	677 - 604	A
85 - 72	B	91 - 84	B	603 - 531	B
71 - 58	C	83 - 76	C	530 - 457	C
57 - 44	D	75 - 68	D	456 - 384	D
< 44	E	< 68	E	< 384	E

IPS S1		IPS S2	
Range	Kelas	Range	Kelas
4.00 - 3.68	A	4.00 - 3.6	A
3.67 - 3.36	B	3.5 - 3.2	B
3.35 - 3.04	C	3.1 - 2.8	C
3.03 - 2.72	D	2.7 - 2.4	D
< 2.72	E	< 2.4	E

Dari belasan variabel ada di atas, dipilih 14 variabel yang diikutkan dalam proses klasifikasi yaitu Akreditasi Perguruan tinggi S1, Program Studi S2, Status Program Studi S1, Bidang Keahlian S2, Sumber Dana S2, IPK S1, IPS Semester 1, IPS Semester 2, Nilai GMAT, Nilai Toefl, Nilai Materi Bidang, Nilai Wawancara, Nilai Skor Akhir, Jarak S1-S2.

3.3 Melakukan Data Mining dengan Menggunakan Algoritma C45

Algoritma C4.5 merupakan salah satu algoritma machine learning. Dengan algoritma ini, mesin (komputer) akan diberikan sekelompok data untuk dipelajari yang disebut learning dataset. Kemudian hasil dari pembelajaran selanjutnya akan digunakan untuk mengolah data-data yang baru yang disebut test dataset . Karena algoritma C4.5 digunakan untuk melakukan klasifikasi, jadi hasil dari pengolahan test dataset berupa pengelompokkan data ke dalam kelas-kelasnya.

Setelah melalui proses mengolah data, dilakukan mining data untuk membentuk suatu pohon keputusan. Pohon dibangun dengan cara membagi data secara rekursif hingga tiap bagian terdiri dari data yang berasal dari kelas yang sama. Bentuk pemecahan (split) yang digunakan untuk membagi data tergantung dari jenis atribut yang digunakan dalam split. Algoritma C4.5 dapat menangani data numerik (kontinyu) dan diskret. Split untuk atribut numerik yaitu mengurutkan contoh berdasarkan atribut kontinyu A, kemudian membentuk minimum permulaan (threshold) M dari contoh-contoh yang ada dari kelas mayoritas pada setiap partisi yang bersebelahan, lalu menggabungkan partisi-partisi yang bersebelahan tersebut dengan kelas mayoritas yang sama. Split untuk atribut diskret A mempunyai bentuk value $(A) \in X$ dimana $X \subset \text{domain}(A)$. Jika suatu set data mempunyai beberapa pengamatan dengan missing value yaitu record dengan beberapa nilai variabel tidak ada, Jika jumlah pengamatan terbatas maka atribut dengan missing value dapat diganti dengan nilai rata-rata dari variabel yang bersangkutan.[Santosa,2007].

Untuk melakukan pemisahan obyek (split) dilakukan tes terhadap atribut dengan mengukur tingkat ketidakmurnian pada sebuah simpul (node). Pada algoritma C.45 menggunakan rasio perolehan (gain ratio). Sebelum menghitung rasio perolehan, perlu menghitung dulu nilai informasi dalam satuan bits dari suatu kumpulan objek. Cara menghitungnya dilakukan dengan menggunakan konsep entropi.

$$E(S) = -p_+ \log_2 p_+ - p_- \log_2 p_-$$

S adalah ruang (data) sampel yang digunakan untuk pelatihan, p_+ adalah jumlah yang bersolusi positif atau mendukung pada data sampel untuk kriteria

tertentu dan p- adalah jumlah yang bersolusi negatif atau tidak mendukung pada data sampel untuk kriteria tertentu. $Entropi(S)$ sama dengan 0, jika semua contoh pada S berada dalam kelas yang sama. $Entropi(S)$ sama dengan 1, jika jumlah contoh positif dan negative dalam S adalah sama. $Entropi(S)$ lebih dari 0 tetapi kurang dari 1, jika jumlah contoh positif dan negative dalam S tidak sama [Mitchell,1997]. Entropi split yang membagi S dengan n record menjadi himpunan-himpunan S_1 dengan n_1 baris dan S_2 dengan n_2 baris adalah :

$$E(S_1, S_2) = \frac{n_1}{n} E(S_1) + \frac{n_2}{n} E(S_2)$$

Kemudian menghitung perolehan informasi dari output data atau variabel dependent y yang dikelompokkan berdasarkan atribut A, dinotasikan dengan $gain(y, A)$. Perolehan informasi, $gain(y, A)$, dari atribut A relative terhadap output data y adalah:

$$Gain(y, A) = entropi(y) - \sum_{c \in nilai(A)} \frac{y_c}{y} entropi(y_c)$$

nilai (A) adalah semua nilai yang mungkin dari atribut A, dan y_c adalah subset dari y dimana A mempunyai nilai c. Term pertama dalam persamaan diatas adalah entropy total y dan term kedua adalah entropy sesudah dilakukan pemisahan data berdasarkan atribut A.

Untuk menghitung rasio perolehan perlu diketahui suatu term baru yang disebut pemisahan informasi (SplitInfo). Pemisahan informasi dihitung dengan cara :

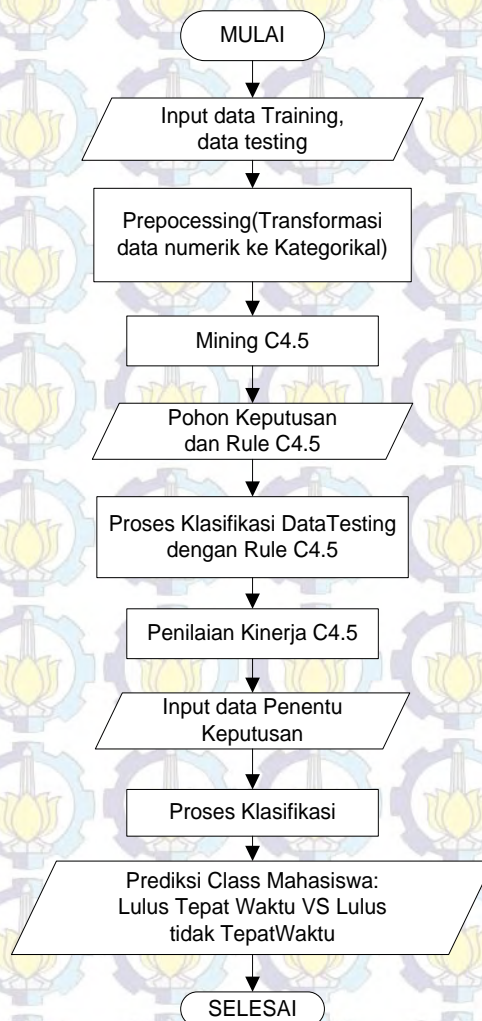
$$SplitInfo(S, A) = - \sum_{i=1}^c \frac{S_i}{S} \log_2 \frac{S_i}{S}$$

bahwa S_1 sampai S_c adalah c subset yang dihasilkan dari pemecahan S dengan menggunakan atribut A yang mempunyai sebanyak c nilai. Selanjutnya rasio perolehan (gain ratio) dihitung dengan cara :

$$GainRatio(S, A) = \frac{Gain(S, A)}{SplitInformation(S, A)}$$

3.4 Membuat Rancangan Flowchart Sistem

Perancangan flowchart akan memudahkan pengembang untuk mengimplementasikan sistem ke dalam bahasa pemrograman, karena akan menjelaskan bagaimana cara kerja sistem dari awal hingga akhir. Pada tesis ini, flowchart berurutan yaitu input data Training dan data Testing, proses preprocessing yaitu transformasi data numeric ke categorical, Mining dengan algoritma C4.5, Pembentukan pohon keputusan dan Rule C4.5, Proses Klasifikasi Data Testing dengan rule C4.5, Penilaian Kinerja C4.5, Input data Penentu Keputusan, Proses Klasifikasi, dan Prediksi Class Mahasiswa : Lulus Tepat Waktu Vs Lulus tidak Tepat Waktu.





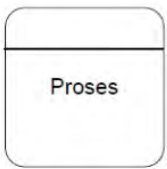
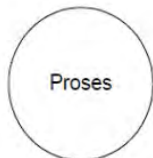
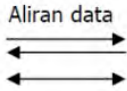
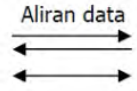
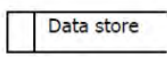

Gambar 3.2
Flowchart Sistem

3.5 Perancangan Basis Data

Diagram Aliran Data / Data Flow Diagram (DFD) adalah sebuah teknis grafis yang menggambarkan aliran informasi dan transformasi yang diaplikasikan saat data bergerak dari input menjadi output. DFD dapat digunakan untuk menyajikan sebuah sistem atau perangkat lunak pada setiap tingkat abstraksi. DFD memberikan suatu mekanisme bagi pemodelan fungsional dan pemodelan informasi.

DFD tingkat 0, disebut juga dengan model sistem fundamental atau model konteks, merepresentasikan seluruh elemen sistem sebagai sebuah lingkaran tunggal dengan data input dan output yang ditunjukkan oleh anak panah yang keluar dan masuk secara berurutan.

Pada penelitian ini, DFD dikelompokkan menjadi dua yakni DFD yang menjelaskan proses yang dilakukan oleh mahasiswa dan proses yang dilakukan oleh seorang dosen. Berikut dengan DFD level 0 untuk sistem pengidentifikasian mahasiswa yang lulus tidak tepat waktu.

Gane/Sarson	Yourdon/De Marco	Keterangan
		Entitas eksternal, dapat berupa orang/unit terkait yang berinteraksi dengan sistem tetapi diluar sistem
		Orang, unit yang mempergunakan atau melakukan transformasi data. Komponen fisik tidak diidentifikasi.
		Aliran data dengan arah khusus dari sumber ke tujuan
		Penyimpanan data atau tempat data direfer oleh proses.

Gambar 3.3
Simbol – Simbol dalam DFD

3.6 Pengujian dan Evaluasi Sistem

Pada tahap ini akan dilakukan pengujian sistem. Pengujian terhadap sistem dilakukan dengan berbagai variasi partisi data sample. Pengujian diukur dengan menggunakan perhitungan recall, precision dan accuracy. Precision adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Sedangkan recall adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. Accuracy didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual. Gambar 3.4 merupakan perumusan recall, precision dan accuracy.

		Nilai sebenarnya	
		TRUE	FALSE
Nilai prediksi	TRUE	TP (True Positive) <i>Correct result</i>	FP (False Positive) <i>Unexpected result</i>
	FALSE	FN (False Negative) <i>Missing result</i>	TN (True Negative) <i>Correct absence of result</i>

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Gambar 3.4
Perhitungan Recall, Precision dan Accuracy

Pada tahap ini akan dilakukan pengujian terhadap sistem yang telah dibuat. Adapun evaluasi berdasarkan pengujian yang dilakukan dengan nilai *precision*, *recall* dan *accuracy* yang dihasilkan dari pengujian sistem yang dilakukan.

3.7 Penarikan Kesimpulan

Setelah dilakukan ujicoba sistem secara keseluruhan maka dapat dilakukan penarikan kesimpulan dari hasil penelitian. Tentunya penelitian ini masih perlu disempurnakan lagi, oleh karena itu perlu adanya saran pengembangan penelitian di masa yang akan datang.

BAB IV

ARSITEKTUR PROGRAM

Pada Bab ini akan dijelaskan detail arsitektur program utama, disertai gambar arsitektur untuk memudahkan pemahaman. Selain itu, akan diberikan analisa mengenai kebutuhan sistem dan spesifikasi dari program. Desain tabel dan desain struktur data juga akan dijelaskan dalam bab ini. Pada desain tabel akan dijelaskan struktur dari semua tabel yang terdapat dalam program Tesis ini. Pada desain struktur data akan dijelaskan struktur semua class yang ada pada program Tesis ini.

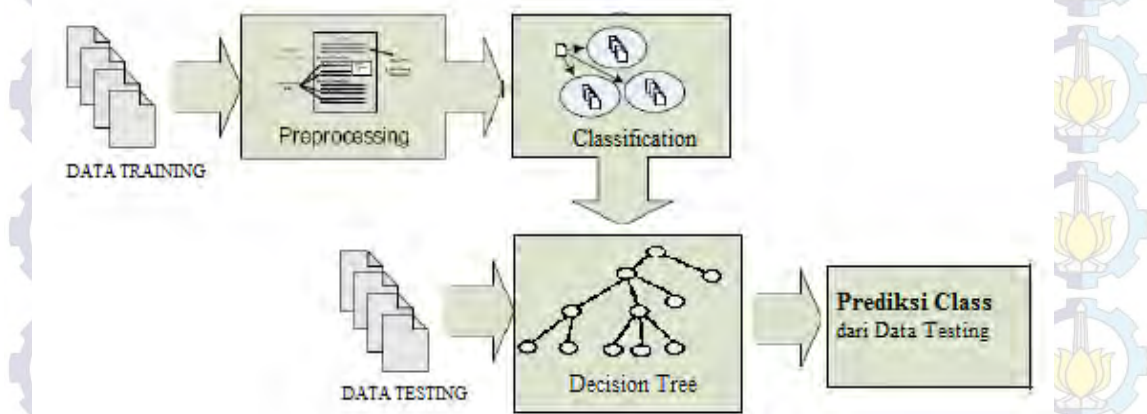
4.1 Gambaran Arsitektur Program Utama

Pada sub bab ini akan dibahas mengenai gambaran detail arsitektur program. Arsitektur program akan dijelaskan secara detail setiap prosesnya. Terdapat 3 proses utama dalam program utama yaitu preprocessing, classification, dan decision tree. Preprocessing adalah proses mengolah data yang ada menjadi data yang sesuai untuk inputan proses selanjutnya. Classification adalah proses pengelompokan data ke dalam kelas-kelas yang ada. Selain itu juga dijelaskan tentang input dan output dari program.

4.1.1 Gambaran Detail Arsitektur

Pada sub bab ini akan dijelaskan tentang gambaran arsitektur program secara detail. Penjelasan arsitektur program akan dibantu dengan gambar. Program utama menerima inputan dari data mahasiswa MMT ITS yang berasal dari sistem akademik dan output adalah prediksi lulus mahasiswa (lulus tepat waktu dan lulus tidak tepat waktu). Pada arsitektur program utama terdapat 3 proses utama yaitu preprocessing, classification, dan decision tree. Ketiga proses ini adalah proses berurutan dan saling berkesinambungan dimana output dari proses sebelum dipakai untuk input untuk proses itu. Pada gambar 4.1 digambarkan detail dari arsitektur program utama.

Pada gambar 4.1 terdapat tiga kotak proses yang disertai tanda panah. Input berupa data mahasiswa angkatan 2009-2012 menuju proses preprocessing. Preprocessing adalah proses mengolah data yang ada menjadi data yang sesuai untuk inputan proses selanjutnya. Proses preprocessing sendiri dibagi menjadi beberapa proses yaitu proses mengisi missing value, dan mengkonversi nilai numeric menjadi huruf dengan skala tertentu.



Gambar 4.1
Arsitektur Program Utama

Setelah proses *preprocessing* adalah proses *classification*. Proses *classification* adalah proses pengelompokan data ke dalam kelas-kelas yang ada. Pada Tesis ini, proses *Classification* menggunakan algoritma C4.5. Input dari proses *classification* adalah *processed data* sedangkan output dari proses *classification* adalah *decision tree*. Selanjutnya *decision tree* berguna untuk memprediksikan class dari data inputan dari dosen ke sistem. Pada Tesis ini terdapat hanya 2 class, yaitu lulus tepat waktu dan lulus tidak tepat waktu. Ketika *Prediksi class* adalah lulus tidak tepat waktu, hal ini dapat dijadikan warning bagi dosen wali untuk lebih memperhatikan mahasiswa tersebut.

4.1.2 Input dan Output program

Pada sub bab ini akan dijelaskan mengenai input dan output program. Input program adalah berupa data Lulusan MMT ITS dari angkatan genap 2009/2010, gasal 2009/2010, genap 2010/2011, gasal 2010/2011, genap

2011/2012, dan gasal 2011/2012 yang telah menyelesaikan studinya di MMT ITS. Data mahasiswa itu didapatkan dari sistem akademik mahasiswa MMT ITS. Data Mahasiswa itu meliputi: NRP, Nama Perguruan Tinggi S1, Jurusan S1, IPK S1, Tahun Lulus S1, Tahun Masuk S2, Nama Perusahaan Tempat Bekerja, Pekerjaan, Nilai GMAT, Nilai Toefl, Nilai Materi Bidang, Nilai Wawancara, Nilai Skor Akhir, Program Studi S2, Bidang Keahlian S2, IPK S2, Sumber dana S2, dan Lama Studi. Contoh dari input program akan dilampirkan pada lampiran A.

Output program adalah prediksi class untuk setiap data yang diinputkan ke sistem. Prediksi class bisa digunakan oleh dosen wali untuk melihat kecenderungan kelulusan mahasiswa tepat waktu atau tidak tepat waktu pada masa awal studinya. Contoh dari output program akan dilampirkan pada lampiran B.

4.2 Analisa Kebutuhan Sistem

Pada sub bab ini akan dijelaskan tentang apa saja kebutuhan tesis ini. Analisa kebutuhan ini meliputi bahasa pemrograman yang dipakai, platform yang dipakai, komponen tambahan yang digunakan, dan program apa saja yang akan dibuat. Adapun setelah dilakukan analisis awal, maka kebutuhan sistem adalah :

- Data training diperoleh dari sistem akademik MMT-ITS. Data ini diperoleh secara manual dengan mengajukan permintaan data ke staff akademik, dan peneliti mendapat data ini berupa softcopy dan hardcopy, lalu peneliti menggabungkan data tersebut.
- Untuk Proses *Preprocessing Data*, Peneliti hanya menggunakan Microsoft excel untuk melakukan *Preprocessing Data*.
- Sebuah interface halaman web utama untuk menjalankan proses klasifikasi menggunakan algoritma C 4.5 dan pruning. Selain itu web ini juga menampilkan proses perhitungan gain dan entropy serta menampilkan rule keputusan yang dibentuk dari proses tersebut.

Proses pengumpulan dan penyeleksian data akan membutuhkan waktu kurang lebih 2 bulan. Proses preprocessing data dan web utama dapat dilakukan secara bertahap setelah proses pengumpulan data selesai. Proses evaluasi akan dilakukan setelah web utama selesai, meliputi perhitungan recall dan precision.

4.3 Desain Tabel dan Struktur Data

Pada Sub bab ini akan dijelaskan mengenai desain tabel dan struktur data yang digunakan. Program ini menggunakan 14 tabel yaitu Tabel Atribut, Tabel Data Keputusan, Tabel Data Keputusan Kinerja, Tabel Data Penentu Keputusan, Tabel Data Training,, Tabel Form Atribut, Tabel Iterasi C45, Tabel Mining C45, Tabel Pohon Keputusan C45, Tabel Rule C45, Tabel Rule Penentu Keputusan, ,Tabel User, Tabel Akreditasi, dan Tabel Program Studi. Tabel Form Atribut digunakan mendefault nilai-nilai atribut pada form pengambilan keputusan. Sedangkan Tabel User digunakan untuk list user dan password. Selain itu dijelaskan mengenai struktur data yang digunakan.

4.3.1 Desain Tabel

Pada Sub bab ini akan menjelaskan mengenai struktur tabel yang dipakai. Tabel dibuat dengan menggunakan MySQL. Ada sebelas tabel yaitu Tabel Atribut, Tabel Data Keputusan, Tabel Data Keputusan Kinerja, Tabel Data Penentu Keputusan, Tabel Data Training, Tabel Form Atribut, Tabel Iterasi C45, Tabel Mining C45, Tabel Pohon Keputusan C45, Tabel Rule C45, Tabel Rule Penentu Keputusan, dan Tabel User. Sembilan Tabel digunakan untuk proses klasifikasi dan pembentukan rule keputusan. Sedangkan 2 tabel lainnya, yaitu Tabel User untuk fungsi login dan Tabel Form Atribut untuk default nilai.

4.3.1.1 Tabel User

Tabel User adalah sebuah tabel yang berisikan daftar username yang dapat login pada website. Adapun field yang dicatat meliputi Id_User, Password, dan Level. Isi detail dari tabel user ditunjukkan pada tabel 4.1.

Tabel 4.1
Struktur Tabel User

No.	Field	Type	Keterangan
1.	Id_User	Varchar(25)	Berisi Username
2.	Password	Varchar(50)	Berisi Password
3.	Level	Varchar(15)	Berisi Level dari User

Tabel User memiliki 3 field yaitu Id_User, Password, dan Level. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel User :

1. Field Id_User adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan Username, bersifat unik dan field ini merupakan Primary Key.
2. Field Password adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan password dari user id yang bersangkutan.
3. Field Level adalah field bertipe data Varchar dengan panjang 15 karakter yang berisikan level dari user id yang bersangkutan. Level dibedakan menjadi admin dan dosen. User level admin dapat mengakses proses training data sampai penentuan keputusan, sedangkan user level dosen hanya dapat mengakses menu penentuan keputusan.

4.3.1.2 Tabel Form Atribut

Tabel Form Atribut adalah digunakan menyimpan atribut dan nilai atribut untuk form atribut (form input data dan form edit data). form ini bersifat statis (tidak digunakan di proses mining). Tabel ini berfungsi untuk mengisi default value untuk atribut-atribut tertentu untuk ditampilkan pada form penentuan keputusan. Adapun field yang dicatat meliputi Id, Atribut, dan Nilai_Atribut. Isi detail dari tabel form atribut ditunjukkan pada tabel 4.2.

Tabel 4.2
Struktur Tabel Form Atribut

No.	Field	Type	Keterangan
1.	Id	Int(3)	Merupakan ID Atribut
2.	Atribut	Varchar(100)	Berisi Nama Atribut
3.	Nilai Atribut	Varchar(100)	Berisi Value dari Atribut

Tabel Form Atribut memiliki 3 field yaitu Id, Atribut, dan Nilai Atribut. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Form Atribut :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Atribut adalah field bertipe data Varchar dengan panjang 100 karakter yang berisikan nama atribut.
3. Field Nilai Atribut adalah field bertipe data Varchar dengan panjang 100 karakter yang berisikan nilai atribut. Nilai-nilai atribut itu digunakan untuk ditampilkan dalam combobox pada form inputan penentu keputusan.

4.3.1.3 Tabel Atribut

Tabel Atribut adalah sebuah tabel yang digunakan utk menyimpan atribut dan nilai atribut guna utk proses mining, dimana setiap iterasi akan ada penghapusan atribut sampai proses perhitungan iterasi selesai. Isi detail dari tabel atribut ditunjukkan pada tabel 4.3.

Tabel 4.3
Struktur Tabel Atribut

No.	Field	Type	Keterangan
1.	Id	Int(3)	Merupakan ID Atribut
2.	Atribut	Varchar(100)	Berisi Nama Atribut
3.	Nilai Atribut	Varchar(100)	Berisi Value dari Atribut

Tabel Form Atribut memiliki 3 field yaitu Id, Atribut, dan Nilai Atribut. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Form Atribut :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Atribut adalah field bertipe data Varchar dengan panjang 100 karakter yang berisikan nama atribut.
3. Field Nilai Atribut adalah field bertipe data Varchar dengan panjang 100 karakter yang berisikan nilai atribut

4.3.1.4 Tabel Data Keputusan

Tabel Data Keputusan adalah sebuah tabel yang menyimpan hasil proses "penentu keputusan". Struktur Tabel data keputusan ditunjukkan pada tabel 4.4.

Tabel 4.4
Struktur Tabel Data Keputusan

No	Field	Type	Keterangan
1.	NRP	Int(15)	Berisi NRP Mahasiswa
2.	Nama PT S1	Varchar(100)	Berisi Nama PT S1
3.	Jurusan S1	Varchar(50)	Berisi Jurusan S1
4.	IPK S1 Numerik	Decimal (5)	Berisi Nilai IPK S1 (numerik)
5.	GMAT Numerik	Int(5)	Berisi Nilai GMAT (numerik)
6.	TOEFL Numerik	Int(5)	Berisi Nilai TOEFL (numerik)
7.	MBid Numerik	Int(5)	Berisi Nilai M.Bidang(numerik)
8.	Wawancara Numerik	Int(5)	Berisi Nilai Wawancara(numerik)
9.	Skor Akhir Numerik	Int(5)	Berisi Skor Akhir (Numerik)
10.	Jarak S1-S2	Varchar(15)	Berisi Jarak S1-S2 (Jauh/Dekat)
11.	IPS Smt1 Numerik	Decimal(5)	Berisi IPS Smt1 Numerik
12.	IPS Smt2 Numerik	Decimal(5)	Berisi IPS Smt2 Numerik
13.	Akreditasi S1	Varchar(5)	Berisi Akreditasi S1
14.	Program Studi	Varchar(5)	Berisi Program Studi S2
15.	Status Program Studi	Varchar(15)	Berisi Status Program Studi
16.	Bidang Keahlian	Varchar(50)	Berisi Bidang Keahlian S2
17.	Biaya	Varchar(50)	Berisi asal biaya S2
18.	IPK S1 Huruf	Varchar (5)	Berisi Nilai IPK S1 (Huruf)
19.	GMAT Huruf	Varchar(5)	Berisi Nilai GMAT (Huruf)
20.	TOEFL Huruf	Varchar(5)	Berisi Nilai TOEFL (Huruf)
21.	MBid Huruf	Varchar(5)	Berisi Nilai M.Bidang(Huruf)
22.	Wawancara Huruf	Varchar(5)	Berisi Nilai Wawancara(Huruf)
23.	Skor Akhir Huruf	Varchar(5)	Berisi Skor Akhir (Huruf)

No	Field	Type	Keterangan
24.	IPS Smter1 Huruf	Varchar(5)	Berisi IPS Smter 1 (Huruf)
25.	IPS Smter2 Huruf	Varchar(5)	Berisi IPS Smter 2 (Huruf)
26.	Keputusan C4.5	Varchar(35)	Berisi Class hasil algoritma C4.5
27.	Id Rule C4.5	Int(5)	Berisi Id Rule keputusan

Tabel Data Keputusan memiliki 27 field yaitu NRP, Nama PT S1, Jurusan S1, Ipk S1 Numerik, GMAT Numerik, TOEFL Numerik, MBid Numerik, Wawancara Numerik, Skor Akhir Numerik, jarak S1 S2, IPS Semester 1 Numerik, IPS Semester 2 Numerik, Akreditasi S1, Program Studi, Status Program Studi, Bidang Keahlian, biaya, Ipk S1 Huruf, GMAT Huruf, TOEFL Huruf, MBid Huruf, Wawancara Huruf, Skor Akhir Huruf, IPS Semester 1 Huruf, IPS Semester 2 Huruf, Keputusan C45, dan Id Rule C45. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Data Keputusan :

1. Field NRP adalah field bertipe data integer dengan panjang 10 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key.
2. Nama PT S1 adalah field bertipe data varchar dengan panjang 100 karakter yang berisikan Nama Perguruan Tinggi S1.
3. Jurusan S1 adalah field bertipe data varchar dengan panjang 50 karakter yang berisikan Jurusan S1
4. Field IPK S1 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPK S1.
5. Field GMAT Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan GMAT.
6. Field TOEFL Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan TOEFL.
7. Field MBid Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan MBid.
8. Field Wawancara Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan Wawancara.

9. Field Skor Akhir Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan Skor Akhir.
10. Field Jarak S1-S2 adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Jarak seorang mahasiswa mengenyam pendidikan S1 dan melanjutkan S2nya. Isi dari Jarak S1-S2 adalah Jauh atau Dekat. Jauh bila jarak tahun lebih besar dari 4 tahun, dan dekat apabila jarak tahun kurang dari sama dengan 4 tahun.
11. Field IPS Semester 1 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 1.
12. Field IPS Semester 2 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 2.
13. Akreditasi S1 adalah Field bertipe data Varchar dengan panjang 5 karakter yang berisikan akreditasi dari Perguruan Tinggi S1
14. Field Program Studi adalah field bertipe data Varchar dengan panjang 15 karakter yang berisikan program studi S2 yang dipilih mahasiswa.
15. Field Status Program Studi adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan apakah jurusan S1 sebidang dengan program studi S2 yang dipilih.
16. Field Bidang Keahlian adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan Bidang Keahlian S2.
17. Field Biaya adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan asal sumber dana pendidikan S2.
18. Field IPK S1 Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan IPK S1 yang telah dikonversi dalam huruf.
19. Field GMAT Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan GMAT yang telah dikonversi dalam huruf.
20. Field TOEFL Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan TOEFL yang telah dikonversi dalam huruf.
21. Field MBid Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan MBid yang telah dikonversi dalam huruf.
22. Field Wawancara Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Wawancara yang telah dikonversi dalam huruf.

23. Field Skor Akhir Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Skor Akhir yang telah dikonversi dalam huruf.
24. Field IPS Semester 1 Huruf adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 1.
25. Field IPS Semester 2 Huruf adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 2.
26. Field Keputusan C4.5 adalah field bertipe data Varchar dengan panjang 35 karakter yang berisikan Class (Lulus Tepat waktu / Lulus Tidak Tepat Waktu) hasil proses klasifikasi menggunakan algoritma C4.5.
27. Field Id Rule C4.5 adalah field bertipe data integer dengan panjang 5 karakter yang berisikan No ID Rule C4.5 yang digunakan untuk proses penentuan keputusan.

4.3.1.5 Tabel Data Keputusan Kinerja

Tabel Data Keputusan Kinerja adalah sebuah tabel yang menyimpan hasil klasifikasi proses "kinerja". Struktur Tabel data keputusan kinerja ditunjukkan pada tabel 4.5.

Tabel 4.5
Struktur Tabel Data Keputusan Kinerja

No	Field	Type	Keterangan
1.	NRP	Int(15)	Berisi NRP Mahasiswa
2.	Nama PT S1	Varchar(100)	Berisi Nama PT S1
3.	Jurusan S1	Varchar(50)	Berisi Jurusan S1
4.	IPK S1 Numerik	Decimal (5)	Berisi Nilai IPK S1 (numerik)
5.	GMAT Numerik	Int(5)	Berisi Nilai GMAT (numerik)
6.	TOEFL Numerik	Int(5)	Berisi Nilai TOEFL (numerik)
7.	MBid Numerik	Int(5)	Berisi Nilai M.Bidang(numerik)
8.	Wawancara Numerik	Int(5)	Berisi Nilai Wawancara(numerik)
9.	Skor Akhir Numerik	Int(5)	Berisi Skor Akhir (Numerik)
10.	Jarak S1-S2	Varchar(15)	Berisi Jarak S1-S2 (Jauh/Dekat)
11.	IPS Smter1 Numerik	Decimal(5)	Berisi IPS Smter 1 Numerik

No	Field	Type	Keterangan
12.	IPS Smt2 Numerik	Decimal(5)	Berisi IPS Smt2 Numerik
13.	Akreditasi S1	Varchar(5)	Berisi Akreditasi S1
14.	Program Studi	Varchar(5)	Berisi Program Studi S2
15.	Status Program Studi	Varchar(15)	Berisi Status Program Studi
16.	Bidang Keahlian	Varchar(50)	Berisi Bidang Keahlian S2
17.	Biaya	Varchar(50)	Berisi asal biaya S2
18.	IPK S1 Huruf	Varchar (5)	Berisi Nilai IPK S1 (Huruf)
19.	GMAT Huruf	Varchar(5)	Berisi Nilai GMAT (Huruf)
20.	TOEFL Huruf	Varchar(5)	Berisi Nilai TOEFL (Huruf)
21.	MBid Huruf	Varchar(5)	Berisi Nilai M.Bidang(Huruf)
22.	Wawancara Huruf	Varchar(5)	Berisi Nilai Wawancara(Huruf)
23.	Skor Akhir Huruf	Varchar(5)	Berisi Skor Akhir (Huruf)
24.	IPS Smt1 Huruf	Varchar(5)	Berisi IPS Smt1 (Huruf)
25.	IPS Smt2 Huruf	Varchar(5)	Berisi IPS Smt2 (Huruf)
26.	Keputusan Asli	Varchar(35)	Berisi Class Sebenarnya
27.	Keputusan C4.5	Varchar(35)	Berisi Class hasil algoritma C4.5
28.	Id Rule C4.5	Int(5)	Berisi Id Rule keputusan

Tabel Data Keputusan Kinerja memiliki 28 field yaitu NRP, Nama PT S1, Jurusan S1, Ipk S1 Numerik, GMAT Numerik, TOEFL Numerik, MBid Numerik, Wawancara Numerik, Skor Akhir Numerik, jarak S1 S2, IPS Semester 1 Numerik, IPS Semester 2 Numerik, Akreditasi S1, Program Studi, Status Program Studi, Bidang Keahlian, biaya, Ipk S1 Huruf, GMAT Huruf, TOEFL Huruf, MBid Huruf, Wawancara Huruf, Skor Akhir Huruf, IPS Semester 1 Huruf, IPS Semester 2 Huruf, Keputusan Asli, Keputusan C45, dan Id Rule C45. Berikut adalah penjelasan field yang terdapat pada Tabel Data Keputusan Kinerja :

1. Field NRP adalah field bertipe data integer dengan panjang 10 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key.
2. Nama PT S1 adalah field bertipe data varchar dengan panjang 100 karakter yang berisikan Nama Perguruan Tinggi S1.

3. Jurusan S1 adalah field bertipe data varchar dengan panjang 50 karakter yang berisikan Jurusan S1
4. Field IPK S1 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPK S1.
5. Field GMAT Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan GMAT.
6. Field TOEFL Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan TOEFL.
7. Field MBid Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan MBid.
8. Field Wawancara Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan Wawancara.
9. Field Skor Akhir Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan Skor Akhir.
10. Field Jarak S1-S2 adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Jarak seorang mahasiswa mengenyam pendidikan S1 dan melanjutkan S2nya. Isi dari Jarak S1-S2 adalah Jauh atau Dekat. Jauh bila jarak tahun lebih besar dari 4 tahun, dan dekat apabila jarak tahun kurang dari sama dengan 4 tahun.
11. Field IPS Semester 1 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 1.
12. Field IPS Semester 2 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 2.
13. Field Akreditasi S1 adalah Field bertipe data Varchar dengan panjang 5 karakter yang berisikan akreditasi dari Perguruan Tinggi S1
14. Field Program Studi adalah field bertipe data Varchar dengan panjang 15 karakter yang berisikan program studi S2 yang dipilih mahasiswa.
15. Field Status Program Studi adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan apakah jurusan S1 sebidang dengan program studi S2 yang dipilih.
16. Field Bidang Keahlian adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan Bidang Keahlian S2.

17. Field Biaya adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan asal sumber dana pendidikan S2.
18. Field IPK S1 Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan IPK S1 yang telah dikonversi dalam huruf.
19. Field GMAT Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan GMAT yang telah dikonversi dalam huruf.
20. Field TOEFL Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan TOEFL yang telah dikonversi dalam huruf.
21. Field MBid Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan MBid yang telah dikonversi dalam huruf.
22. Field Wawancara Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Wawancara yang telah dikonversi dalam huruf.
23. Field Skor Akhir Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Skor Akhir yang telah dikonversi dalam huruf.
24. Field IPS Semester 1 Huruf adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 1.
25. Field IPS Semester 2 Huruf adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 2.
26. Field Keputusan Asli adalah field bertipe data Varchar panjang 35 karakter yang berisikan Class (Lulus Tepat waktu / Lulus Tidak Tepat Waktu) yang sebenarnya.
27. Field Keputusan C4.5 adalah field bertipe data Varchar dengan panjang 35 karakter yang berisikan Class (Lulus Tepat waktu / Lulus Tidak Tepat Waktu) hasil proses klasifikasi menggunakan algoritma C4.5.
28. Field Id Rule C4.5 adalah field bertipe data integer dengan panjang 5 karakter yang berisikan No ID Rule C4.5 yang digunakan untuk proses penentuan keputusan.

4.3.1.6 Tabel Data Penentu Keputusan

Tabel Data Penentu Keputusan adalah sebuah tabel yang digunakan menyimpan atribut dan nilai atribut sementara yang akan digunakan untuk proses

"penentu keputusan" dan proses "kinerja". Adapun field yang dicatat meliputi Id, Atribut, dan Nilai_Atribut. Isi detail dari tabel atribut ditunjukkan pada tabel 3.6.

Tabel 4.6
Struktur Tabel Data Penentu Keputusan

No.	Field	Type	Keterangan
1.	Id	Int(3)	Merupakan ID Atribut
2.	Atribut	Varchar(100)	Berisi Nama Atribut
3.	Nilai Atribut	Varchar(100)	Berisi Value dari Atribut

Tabel Form Data Penentu Keputusan memiliki 3 field yaitu Id, Atribut, dan Nilai Atribut. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Form Data Penentu Keputusan :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Atribut adalah field bertipe data Varchar dengan panjang 100 karakter yang berisikan nama atribut.
3. Field Nilai Atribut adalah field bertipe data Varchar dengan panjang 100 karakter yang berisikan nilai atribut

4.3.1.7 Tabel Data Training

Tabel Data Training adalah sebuah tabel yang menyimpan data training. Struktur tabel data training ditunjukkan pada tabel 4.7.

Tabel 4.7
Struktur Tabel Data Training

No	Field	Type	Keterangan
1.	NRP	Int(15)	Berisi NRP Mahasiswa
2.	Nama PT S1	Varchar(100)	Berisi Nama PT S1
3.	Jurusan S1	Varchar(50)	Berisi Jurusan S1
4.	IPK S1 Numerik	Decimal (5)	Berisi Nilai IPK S1 (numerik)

No	Field	Type	Keterangan
5.	GMAT Numerik	Int(5)	Berisi Nilai GMAT (numerik)
6.	TOEFL Numerik	Int(5)	Berisi Nilai TOEFL (numerik)
7.	MBid Numerik	Int(5)	Berisi Nilai M.Bidang(numerik)
8.	Wawancara Numerik	Int(5)	Berisi Nilai Wawancara(numerik)
9.	Skor Akhir Numerik	Int(5)	Berisi Skor Akhir (Numerik)
10.	Jarak S1-S2	Varchar(15)	Berisi Jarak S1-S2 (Jauh/Dekat)
11.	IPS Smter1 Numerik	Decimal(5)	Berisi IPS Smter 1 Numerik
12.	IPS Smter2 Numerik	Decimal(5)	Berisi IPS Smter 2 Numerik
13.	Akreditasi S1	Varchar(5)	Berisi Akreditasi S1
14.	Program Studi	Varchar(5)	Berisi Program Studi S2
15.	Status Program Studi	Varchar(15)	Berisi Status Program Studi
16.	Bidang Keahlian	Varchar(50)	Berisi Bidang Keahlian S2
17.	Biaya	Varchar(50)	Berisi asal biaya S2
18.	IPK S1 Huruf	Varchar (5)	Berisi Nilai IPK S1 (Huruf)
19.	GMAT Huruf	Varchar(5)	Berisi Nilai GMAT (Huruf)
20.	TOEFL Huruf	Varchar(5)	Berisi Nilai TOEFL (Huruf)
21.	MBid Huruf	Varchar(5)	Berisi Nilai M.Bidang(Huruf)
22.	Wawancara Huruf	Varchar(5)	Berisi Nilai Wawancara(Huruf)
23.	Skor Akhir Huruf	Varchar(5)	Berisi Skor Akhir (Huruf)
24.	IPS Smter1 Huruf	Varchar(5)	Berisi IPS Smter 1 (Huruf)
25.	IPS Smter2 Huruf	Varchar(5)	Berisi IPS Smter 2 (Huruf)
26.	Class	Varchar(35)	Berisi Class Sebenarnya
27.	Status Data	Varchar(35)	Berisi Status Data (Training dan Testing)

Tabel Data Training memiliki 27 field yaitu NRP, Nama PT S1, Jurusan S1, Ipk S1 Numerik, GMAT Numerik, TOEFL Numerik, MBid Numerik, Wawancara Numerik, Skor Akhir Numerik, jarak S1 S2, IPS Semester 1 Numerik, IPS Semester 2 Numerik, Akreditasi S1, Program Studi, Status Program Studi, Bidang Keahlian, biaya, Ipk S1 Huruf, GMAT Huruf, TOEFL Huruf, MBid

Huruf, Wawancara Huruf, Skor Akhir Huruf, IPS Semester 1 Huruf, IPS Semester 2 Huruf, Keputusan Asli, Class, dan Status Data . Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Data Training :

1. Field NRP adalah field bertipe data integer dengan panjang 10 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key.
2. Nama PT S1 adalah field bertipe data varchar dengan panjang 100 karakter yang berisikan Nama Perguruan Tinggi S1.
3. Jurusan S1 adalah field bertipe data varchar dengan panjang 50 karakter yang berisikan Jurusan S1
4. Field IPK S1 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPK S1.
5. Field GMAT Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan GMAT.
6. Field TOEFL Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan TOEFL.
7. Field MBid Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan MBid.
8. Field Wawancara Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan Wawancara.
9. Field Skor Akhir Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan Skor Akhir.
10. Field Jarak S1-S2 adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Jarak seorang mahasiswa mengenyam pendidikan S1 dan melanjutkan S2nya. Isi dari Jarak S1-S2 adalah Jauh atau Dekat. Jauh bila jarak tahun lebih besar dari 4 tahun, dan dekat apabila jarak tahun kurang dari sama dengan 4 tahun.
11. Field IPS Semester 1 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 1.
12. Field IPS Semester 2 Numerik adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 2.
13. Akreditasi S1 adalah Field bertipe dataVarchar dengan panjang 5 ka rakter yang berisikan akreditasi dari Perguruan Tinggi S1

14. Field Program Studi adalah field bertipe data Varchar dengan panjang 15 karakter yang berisikan program studi S2 yang dipilih mahasiswa.
15. Field Status Program Studi adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan apakah jurusan S1 sebidang dengan program studi S2 yang dipilih.
16. Field Bidang Keahlian adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan Bidang Keahlian S2.
17. Field Biaya adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan asal sumber dana pendidikan S2.
18. Field IPK S1 Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan IPK S1 yang telah dikonversi dalam huruf.
19. Field GMAT Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan GMAT yang telah dikonversi dalam huruf.
20. Field TOEFL Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan TOEFL yang telah dikonversi dalam huruf.
21. Field MBid Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan MBid yang telah dikonversi dalam huruf.
22. Field Wawancara Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Wawancara yang telah dikonversi dalam huruf.
23. Field Skor Akhir Huruf adalah field bertipe data Varchar dengan panjang 5 karakter yang berisikan Skor Akhir yang telah dikonversi dalam huruf.
24. Field IPS Semester 1 Huruf adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 1.
25. Field IPS Semester 2 Huruf adalah field bertipe data Int dengan panjang 5 karakter yang berisikan IPS Semester 2.
26. Field Class adalah field bertipe data Varchar dengan panjang 35 karakter yang berisikan Class (Lulus Tepat waktu / Lulus Tidak Tepat Waktu) yang sebenarnya.
27. Field Status Data adalah field bertipe data Varchar dengan panjang 35 karakter yang berisikan status data yaitu data Training.

4.3.1.8 Tabel Iterasi C45

Tabel Iterasi C45 adalah sebuah tabel yang digunakan menyimpan seluruh perhitungan iterasi awal hingga akhir yang berasal dari tabel mining_c45. Adapun field yang dicatat meliputi Id, Iterasi, Atribut Gain Ratio Max, Atribut, Nilai Atribut, jumlah kasus total, jumlah kasus tidak tepat waktu, jumlah kasus tepat waktu. Entropy, Informasi Gain, Split Info, dan Gain Ratio. Isi detail dari tabel Iterasi C45 ditunjukkan pada tabel 4.8.

Tabel 4.8
Struktur Tabel Iterasi C45

No.	Field	Type	Keterangan
1.	Id	Int(11)	Merupakan ID Atribut
2.	Iterasi	Varchar(3)	Berisi nomor iterasi
3.	Atribut Gain Ratio Max	Varchar(50)	Berisikan Nama Atribut dengan gain max
4.	Atribut	Varchar(50)	Berisikan Nama Atribut
5.	Nilai Atribut	Varchar(50)	Berisikan Nilai Atribut
6.	jumlah kasus total	Varchar(10)	Berisikan Total kasus
7.	jumlah kasus tidak tepat waktu	Varchar(10)	Berisikan total kasus yang tidak tepat waktu
8.	jumlah kasus tepat waktu	Varchar(10)	Berisikan total kasus yang tepat waktu
9.	Entropy	Varchar(10)	Berisikan Nilai Entropy
10.	Informasi Gain	Varchar(10)	Berisikan Nilai Informasi Gain
11.	Split Info	Varchar(10)	Berisikan Nilai Split Info
12.	Gain Ratio	Varchar(10)	Berisikan Gain Ratio

Tabel Form Iterasi C45 memiliki 12 field yaitu Id, Iterasi, Atribut Gain Ratio Max, Atribut, Nilai Atribut, jumlah kasus total, jumlah kasus tidak tepat waktu, jumlah kasus tepat waktu. Entropy, Informasi Gain, Split Info, dan Gain Ratio. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Iterasi C45 :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Iterasi adalah field bertipe data Varchar dengan panjang 3 karakter yang berisikan nomor iterasi.
3. Field Atribut Gain Ratio Max adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan nilai atribut yang memiliki Gain Ratio Maximum.
4. Field Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan nama atribut.
5. Field Nilai Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan nilai atribut.
6. Field jumlah kasus total adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan Jumlah kasus total dari atribut dengan nilai atribut bersangkutan.
7. Field jumlah kasus tidak tepat waktu adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan jumlah kasus tidak tepat waktu dari atribut dengan nilai atribut bersangkutan.
8. Field jumlah kasus tepat waktu adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan jumlah kasus tepat waktu dari atribut dengan nilai atribut bersangkutan.
9. Field Entropy adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai entropy.
10. Field Informasi Gain adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Informasi Gain.
11. Field Split Info adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Split Info.
12. Field Gain Ratio adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Gain Ratio.

4.3.1.9 Tabel Mining C45

Tabel Mining C45 adalah sebuah tabel yang digunakan menyimpan proses perhitungan entropy, inf gain, split info, dan gain ratio. Data yang disimpan bersifat sementara karena tiap proses iterasi tabel ini dikosongkan. Adapun field yang dicatat meliputi Id, Atribut, Nilai Atribut, jumlah kasus total, jumlah kasus tidak tepat waktu, jumlah kasus tepat waktu. Entropy, Informasi Gain, Informasi Gain temp, Split Info, Split Info temp, dan Gain Ratio. Isi detail dari tabel Mining C45 ditunjukkan pada tabel 4.9.

Tabel 4.9
Struktur Tabel Mining C45

No.	Field	Type	Keterangan
1.	Id	Int(11)	Merupakan ID Atribut
2.	Atribut	Varchar(50)	Berisikan Nama Atribut
3.	Nilai Atribut	Varchar(50)	Berisikan Nilai Atribut
4.	jumlah kasus total	Varchar(10)	Berisikan Total kasus
5.	jumlah kasus tidak tepat waktu	Varchar(10)	Berisikan total kasus yang tidak tepat waktu
6.	jumlah kasus tepat waktu	Varchar(10)	Berisikan total kasus yang tepat waktu
7.	Entropy	Varchar(10)	Berisikan Nilai Entropy
8.	Informasi Gain	Varchar(10)	Berisikan Nilai Informasi Gain
9.	Informasi Gain Temp	Varchar(10)	Berisikan Nilai Informasi Gain Temporary
10.	Split Info	Varchar(10)	Berisikan Nilai Split Info
11.	Split Info Temp	Varchar(10)	Berisikan Nilai Split Info Temporary
12.	Gain Ratio	Varchar(10)	Berisikan Gain Ratio

Tabel Form Mining C45 memiliki 12 field yaitu Id, Atribut, Nilai Atribut, jumlah kasus total, jumlah kasus tidak tepat waktu, jumlah kasus tepat waktu. Entropy, Informasi Gain, Informasi Gain temp, Split Info, Split Info temp, dan

Gain Ratio. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Iterasi C45 :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan nama atribut.
3. Field Nilai Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan nilai atribut.
4. Field jumlah kasus total adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan Jumlah kasus total dari atribut dengan nilai atribut bersangkutan.
5. Field jumlah kasus tidak tepat waktu adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan jumlah kasus tidak tepat waktu dari atribut dengan nilai atribut bersangkutan.
6. Field jumlah kasus tepat waktu adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan jumlah kasus tepat waktu dari atribut dengan nilai atribut bersangkutan.
7. Field Entropy adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai entropy.
8. Field Informasi Gain adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Informasi Gain.
9. Field Informasi Gain Temp adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Informasi Gain Temporary.
10. Field Split Info adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Split Info.
11. Field Split Info adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Split Info Temporary.
12. Field Gain Ratio adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai Gain Ratio.

4.3.1.10 Tabel Pohon Keputusan C45

Tabel Pohon Keputusan C45 adalah sebuah tabel yang digunakan menyimpan atribut root dan subtree yg diperoleh dari mining. Adapun field yang dicatat meliputi Id, Atribut, Nilai Atribut, id Parent, jumlah kasus tidak tepat waktu, jumlah kasus tepat waktu. Keputusan, diproses, kondisi atribut, dan looping kondisi. Isi detail dari tabel Pohon Keputusan C45 ditunjukkan pada tabel 3.10.

Tabel 4.10
Struktur Tabel Pohon Keputusan C45

No.	Field	Type	Keterangan
1.	Id	Int(11)	Merupakan ID Atribut
2.	Atribut	Varchar(50)	Berisikan Nama Atribut
3.	Id Parent	Varchar(3)	Berisikan ID Parent
4.	Nilai Atribut	Varchar(50)	Berisikan Nilai Atribut
5.	jumlah kasus tidak tepat waktu	Varchar(10)	Berisikan total kasus yang tidak tepat waktu
6.	jumlah kasus tepat waktu	Varchar(10)	Berisikan total kasus yang tepat waktu
7.	Keputusan	Varchar(25)	Berisikan Keputusan
8.	diproses	Varchar(10)	Berisikan flag sudah diproses
9.	kondisi atribut	Varchar(255)	Berisikan Kondisi Atribut
10.	looping kondisi	Varchar(15)	Berisikan looping Kondisi

Tabel Form Pohon Keputusan C45 memiliki 10 field yaitu Id, Atribut, Nilai Atribut, id Parent, jumlah kasus tidak tepat waktu, jumlah kasus tepat waktu. Keputusan, diproses, kondisi atribut, dan looping kondisi. Berikut adalah penjelasan masing-masing field yang terdapat pada Tabel Iterasi C45 :

1. Field Id adalah field bertipe data integer dengan panjang 3 ka rakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan nama atribut.

3. Field Nilai Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan nilai atribut.
4. Field Id_Parent adalah field bertipe data Varchar dengan panjang 3 karakter yang berisikan Id Parent.
5. Field jumlah kasus tidak tepat waktu adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan jumlah kasus tidak tepat waktu dari atribut dengan nilai atribut bersangkutan.
6. Field jumlah kasus tepat waktu adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan jumlah kasus tepat waktu dari atribut dengan nilai atribut bersangkutan.
7. Field Keputusan adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan Class (Lulus Tepat waktu / Lulus Tidak Tepat Waktu) yang sebenarnya.
8. Field Diproses adalah field bertipe data Varchar dengan panjang 10 karakter yang berisikan nilai apakah data tersebut telah diproses atau belum.
9. Field Kondisi Atribut adalah field bertipe data Varchar dengan panjang 255 karakter yang berisikan kondisi atribut.
10. Field looping atribut adalah field bertipe data Varchar dengan panjang 15 karakter yang berisikan nilai apakah data itu telah di looping atau belum.

4.3.1.11 Tabel Rule C45

Tabel Rule C45 adalah sebuah tabel yang digunakan menyimpan rule hasil generate pohon_keputusan_c45. Adapun field yang dicatat meliputi Id, Id Parent, Rule dan Keputusan. Isi detail dari tabel Rule C45 ditunjukkan pada tabel 4.11.

Tabel 4.11
Struktur Tabel Rule C45

No.	Field	Type	Keterangan
1.	Id	Int(3)	Merupakan ID Rule
2.	Id Parent	Int(3)	Berisi Id Parent Rule
3.	Rule	Varchar(255)	Berisi value dari Rule
4.	Keputusan	Varchar(25)	Berisi Class Keputusan

Table Rule C45 memiliki 4 field yaitu Id, Id Parent, Rule dan Keputusan. Berikut adalah penjelasan masing-masing field pada Tabel Rule C45 :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Id Parent adalah field bertipe data integer dengan panjang 3 karakter yang berisikan Id Parent dari Rule yang bersangkutan.
3. Field Rule adalah field bertipe data Varchar dengan panjang 255 karakter yang berisikan value Rule.
4. Field Keputusan adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan Class (Lulus Tepat waktu / Lulus Tidak Tepat Waktu) dari hasil proses klasifikasi.

4.3.1.12 Tabel Rule Penentu Keputusan

Tabel Rule Penentu Keputusan adalah sebuah tabel yang berisi split (potongan) atribut dan nilai atribut hasil generate rule yang digunakan untuk proses "penentu keputusan" dan proses "kinerja". Adapun field yang dicatat meliputi Id, Id Rule, Atribut, Nilai Atribut, Keputusan, dan Cocok. Isi detail dari tabel Rule Penentu Keputusan ditunjukkan pada tabel 4.12.

Tabel 4.12
Struktur Tabel Rule Penentu Keputusan

No.	Field	Type	Keterangan
1.	Id	Int(3)	Merupakan ID Rule
2.	Id Rule	Int(3)	Berisi Id Parent Rule
3.	Atribut	Varchar(50)	Berisi Nama Atribut
4.	Nilai Atribut	Varchar(50)	Berisi Nilai dari Atribut
5.	Keputusan	Varchar (25)	Berisi Class Keputusan
6.	Cocok	Varchar(15)	Berisi Kecocokan keputusan

Table Rule Penentu Keputusan memiliki 6 field yaitu , Id Rule, Atribut, Nilai Atribut, Keputusan, dan Cocok. Berikut adalah penjelasan masing-masing field yang ditunjukkan pada Tabel Rule penentu Keputusan :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Field Id Rule adalah field bertipe data integer dengan panjang 3 karakter yang berisikan Id Rule.
3. Field Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan Nama Atribut.
4. Field Nilai Atribut adalah field bertipe data Varchar dengan panjang 50 karakter yang berisikan Nilai Atribut.
5. Field Keputusan adalah field bertipe data Varchar dengan panjang 25 karakter yang berisikan Class (Lulus Tepat waktu / Lulus Tidak Tepat Waktu) dari hasil proses klasifikasi.
6. Field Cocok adalah field bertipe data Varchar dengan panjang 15 karakter yang berisikan Nilai apakah hasil cocok dengan keputusan asli.

4.3.1.13 Tabel Akreditasi

Tabel Akreditasi adalah sebuah tabel yang berisi list akreditasi dari sebuah program studi pada universitas yang terkait. Adapun field yang dicatat meliputi Id, Nama PT S1, Jurusan S1, dan Akreditasi. Isi detail dari tabel Akreditasi ditunjukkan pada tabel 4.13.

Tabel 4.13
Struktur Tabel Akreditasi

No.	Field	Type	Keterangan
1.	Id	Int(3)	Merupakan ID Rule
2.	Nama PT S1	Varchar(100)	Berisi Nama PT S1
3.	Jurusan S1	Varchar(50)	Berisi Jurusan S1
4.	Akreditasi S1	Varchar(5)	Berisi Akreditasi S1

Table Akreditasi memiliki 4 field yaitu , Id, Nama PT S1, Jurusan S1, dan Akreditasi. Berikut adalah penjelasan masing-masing field yang ditunjukkan pada Tabel Akreditasi :

1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Nama PT S1 adalah field bertipe data varchar dengan panjang 100 karakter yang berisikan Nama Perguruan Tinggi S1.
3. Jurusan S1 adalah field bertipe data varchar dengan panjang 50 karakter yang berisikan Jurusan S1
4. Field Akreditasi S1 adalah Field bertipe data Varchar dengan panjang 5 karakter yang berisikan akreditasi dari Perguruan Tinggi S1.

4.3.1.14 Tabel Program Studi

Tabel Akreditasi adalah sebuah tabel yang berisi list akreditasi dari sebuah program studi pada universitas yang terkait. Adapun field yang dicatat meliputi Id, Jurusan S1, program studi dan Status Program Studi. Isi detail dari tabel Program Studi ditunjukkan pada tabel 4.14.

Tabel 4.14
Struktur Tabel Program Studi

No.	Field	Type	Keterangan
1.	Id	Int(3)	Merupakan ID Rule
2.	Jurusan S1	Varchar(50)	Berisi Jurusan S1
3.	Program Studi	Varchar(5)	Berisi Program Studi S2
4.	Status Program Studi	Varchar(15)	Berisi Status Program Studi

Table Program Studi memiliki 4 field yaitu Id, Jurusan S1, program studi dan Status Program Studi. Berikut adalah penjelasan masing-masing field yang ditunjukkan pada Tabel Program Studi :

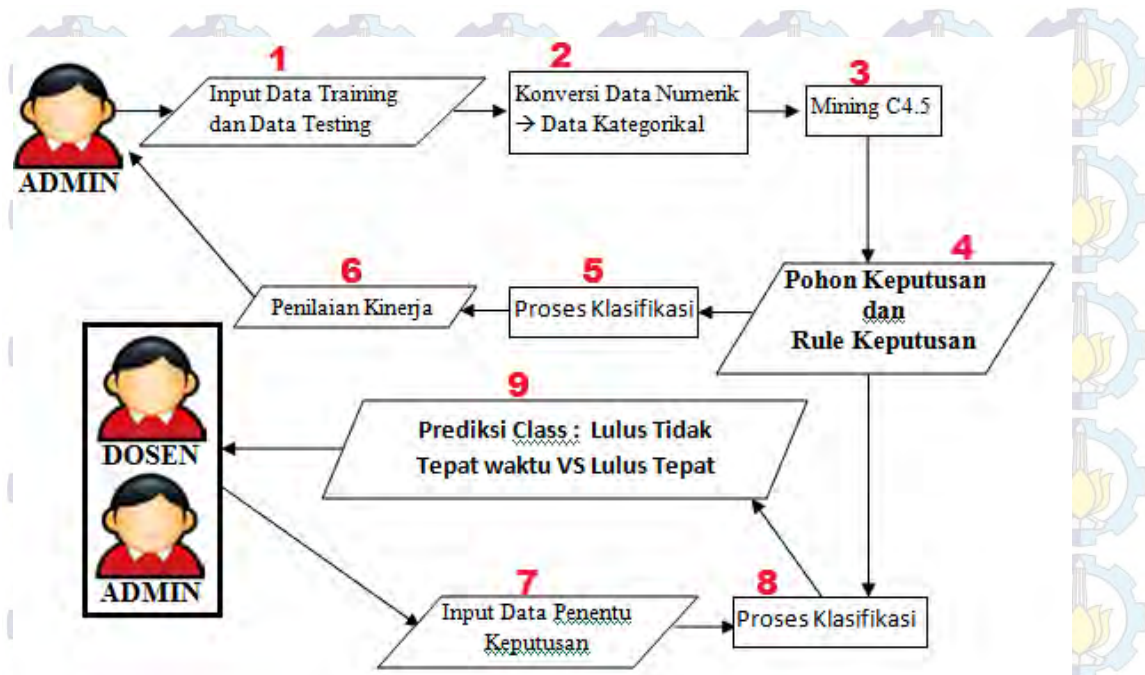
1. Field Id adalah field bertipe data integer dengan panjang 3 karakter yang berisikan id atribut, bersifat unik dan field ini merupakan Primary Key, auto increment.
2. Jurusan S1 adalah field bertipe data varchar dengan panjang 50 karakter yang berisikan Jurusan S1
3. Field Program Studi adalah field bertipe data Varchar dengan panjang 15 karakter yang berisikan program studi S2 yang dipilih mahasiswa.
4. Field Status Program Studi adalah field bertipe data Varchar dengan 25 karakter, berisikan jurusan S1 sebidang dengan program studi S2 yang dipilih

4.3.2 Desain Struktur Data

Pada Sub Bab ini akan dijelaskan tentang *procedure* yang digunakan pada program Tesis ini. *Procedure* - *procedure* tersebut digunakan dalam proses yang ada pada arsitektur program. Terdapat beberapa *procedure* dalam program Tesis ini meliputi *procedure* perhitunganC45, *procedure* perhitungan prepruning dan *procedure* PenentuKeputusan. *Procedure* PerhitunganC45 adalah *procedure* untuk melakukan inialisasi atribut, perhitungan gain, perhitungan entropy dan perhitungan split info. *Procedure* PerhitunganPrepruning adalah *procedure* untuk menentukan apakah cabang tersebut dipruning atau tidak. *Procedure* PenentuanKeputusan adalah *procedure* untuk menentukan class dari data yang dicobakan pada sistem. Implementasi program dari *procedure* – *procedure* itu akan dijelaskan pada bab 5.1.1 Implementasi Program.

4.4 Alur Kerja

Pada sub bab ini akan dijelaskan mengenai langkah-langkah penyelesaian Tesis ini. Alur kerja meliputi Proses input data training dan data testing, Proses konversi Data Numerik menjadi data Kategorikal, Proses Mining C4.5, Pembentukan Pohon keputusan dan Rule Keputusan, Penilaian Kinerja, Input data penentuan keputusan, proses klasifikasi, Prediksi Class : Lulus tidak tepat waktu atau Lulus tepat waktu. Pada Tesis ini User dibedakan menjadi 2 level yaitu admin dan dosen. Dosen hanya bisa menginputkan data penentu keputusan ke sistem, sedangkan admin mempunyai otoritas untuk menginputkan data training.



Gambar 4.2
Alur Kerja

Pertama – tama admin, menginputkan data training dan data testing ke dalam sistem , gambar 4.2 proses nomor 1. Setelah itu, sistem akan melakukan konversi pada data numerik menjadi data kategorikal, gambar 4.2 proses nomor 2. Kemudian dilanjutkan dengan proses mining dengan algoritma C4.5 sehingga terbentuklah pohon keputusan dan rule keputusan, gambar 4.2 proses nomor 3 dan 4. Pohon keputusan dan rule keputusan akan digunakan proses klasifikasi data testing nomor 5. Selanjutnya proses penilaian kinerja pada data testing (gambar 4.2 proses nomor 6). User admin dapat melihat tabel penilaian, pohon keputusan dan rule keputusan.

User Admin dan User Dosen, bisa menginputkan data pada menu penentuan keputusan. Proses ini digambarkan pada gambar 4.2 proses nomor 7,8, dan 9. Pada proses ini Output berupa Prediksi Class : Lulus tidak Tepat Waktu dan Lulus Tepat Waktu. Proses ini menggunakan pohon keputusan dan rule keputusan pada proses sebelumnya yang dijalankan oleh user admin.

BAB V

IMPLEMENTASI DAN UJI COBA SISTEM

Pada bab ini akan dibahas mengenai implementasi program uji coba sistem. Implementasi sistem meliputi implementasi program dan implementasi antar muka. Implementasi program meliputi penjelasan sub program per bagian, sedangkan untuk implementasi antar muka menjelaskan tentang fungsi tampilan website secara keseluruhan. Pada sub bab Ujicoba Sistem akan dijelaskan ujicoba sistem dengan berbagai skenario beserta penjelasannya.

5.1 Implementasi Sistem

Pada sub bab ini akan dibahas mengenai seluk beluk implementasi sistem. Implementasi sistem yang dibahas meliputi implementasi program dan implementasi antar muka perangkat Lunak. Implementasi program adalah implementasi source code yang ada pada sistem. Sedangkan implementasi antar muka perangkat lunak adalah Implementasi fungsi tampilan website.

5.1.1 Implementasi Program

Implementasi program yang dijelaskan meliputi implementasi program pada klasifikasi data menggunakan algoritma C45, implementasi program prepruning, implementasi program perhitungan kinerja, dan implementasi program penentuan keputusan. Implementasi program klasifikasi data menggunakan algoritma C45 terbagi menjadi sub function, yaitu perhitunganC45, insertAtributPohonKeputusan, dan getInfGainMax. Function perhitunganC45 terdiri dari 4 proses yaitu perhitungan entropy, perhitungan information gain, perhitungan split info dan perhitungan gain ratio. Perhitungan entropy dijelaskan pada Segmen Program 5.1, perhitungan informasi gain dijelaskan pada Segmen Program 5.2, perhitungan split info dijelaskan pada Segmen Program 5.3 dan perhitungan gain ratio dijelaskan pada Segmen Program 5.4.

Segmen Program 5.1 Function perhitunganC45 Part I

```
1: function perhitunganC45($atribut, $nilai_atribut){
2: if ($getJumlahKasusTotalEntropy == 0 OR
   $getJumlahKasusTepatWaktuEntropy == 0 OR
   $getJumlahKasusTidakTepatWaktuEntropy == 0) {
3:   $getEntropy = 0;
4: } else if ($getJumlahKasusTepatWaktuEntropy ==
   $getJumlahKasusTidakTepatWaktuEntropy) {
5:   $getEntropy = 1;
6: } else {
7:   $perbandingan_tepat_waktu = $getJumlahKasusTepatWaktuEntropy
   / $getJumlahKasusTotalEntropy;
8:   $perbandingan_tidak_tepat_waktu =
   $getJumlahKasusTidakTepatWaktuEntropy /
   $getJumlahKasusTotalEntropy;
9:   $rumusEntropy = (-( $perbandingan_tepat_waktu ) *
   log( $perbandingan_tepat_waktu, 2 ) +
   ( - ( $perbandingan_tidak_tepat_waktu ) *
   log( $perbandingan_tidak_tepat_waktu, 2 ) ) );
10:   $getEntropy = round( $rumusEntropy, 4 ); }
11: mysql_query("UPDATE mining_c45 SET entropy = $getEntropy
   WHERE id = $idEntropy");}
```

Segmen Program 5.1 merupakan potongan program untuk proses perhitungan entropy dan proses update nilai entropy pada database

Segmen Program 5.2 Function perhitunganC45 Part II

```
1: function perhitunganC45($atribut, $nilai_atribut){
2: $sqlJumlahKasusTotalInfGain = mysql_query("SELECT
   jml_kasus_total, entropy FROM mining_c45 WHERE atribut =
   'Total'");
3: $rowJumlahKasusTotalInfGain =
   mysql_fetch_array($sqlJumlahKasusTotalInfGain);
4: $getJumlahKasusTotalInfGain =
   $rowJumlahKasusTotalInfGain['jml_kasus_total'];
5: $getInfGain = (-( $getJumlahKasusTotalEntropy /
   $getJumlahKasusTotalInfGain ) * ( $getEntropy ) );
6: mysql_query("UPDATE mining_c45 SET inf_gain_temp = $getInfGain
   WHERE id = $idEntropy");
7: $getEntropy = $rowJumlahKasusTotalInfGain['entropy'];
8: $sqlAtributInfGain = mysql_query("SELECT SUM(inf_gain_temp) as
   inf_gain FROM mining_c45 WHERE atribut = '$getAtribut'");
9: while ($rowAtributInfGain =
   mysql_fetch_array($sqlAtributInfGain)) {
10:   $getAtributInfGain = $rowAtributInfGain['inf_gain'];
11:   $getInfGainFix = round(( $getEntropy +
   $getAtributInfGain ), 4 );
12:   mysql_query("UPDATE mining_c45 SET inf_gain =
   $getInfGainFix WHERE atribut = '$getAtribut'");}
```

Segmen Program 5.2 merupakan potongan program untuk proses perhitungan informasi Gain. Pada perhitungan ini terdapat looping untuk

perhitungan information gain setiap atribut dan proses update nilai gain pada database.

Segmen Program 5.3 Function perhitunganC45 Part III

```
1: function perhitunganC45($atribut, $nilai_atribut){
2: $getSplitInfo = (($getJumlahKasusTotalEntropy /
  $getJumlahKasusTotalInfGain) *
  (log(($getJumlahKasusTotalEntropy /
  $getJumlahKasusTotalInfGain),2)));
3: mysql_query("UPDATE mining_c45 SET split_info_temp =
  $getSplitInfo WHERE id = $idEntropy");
4: $sqlAtributSplitInfo = mysql_query("SELECT SUM(split_info_temp)
  as split_info FROM mining_c45 WHERE atribut = '$getAtribut'");
5: while ($rowAtributSplitInfo =
  mysql_fetch_array($sqlAtributSplitInfo)){
6: $getAtributSplitInfo = $rowAtributSplitInfo['split_info'];
7: $getSplitInfoFix = -(round($getAtributSplitInfo,4));
8: mysql_query("UPDATE mining_c45 SET split_info =
  $getSplitInfoFix WHERE atribut = '$getAtribut'");}}
```

Segmen Program 5.3 merupakan potongan program untuk proses perhitungan split info. Pada perhitungan ini terdapat looping untuk perhitungan split info setiap atribut dan proses update nilai split info pada database.

Segmen Program 5.4 Function perhitunganC45 Part IV

```
1: function perhitunganC45($atribut, $nilai_atribut){
2: $sqlGainRatio = mysql_query("SELECT id, inf_gain, split_info
  FROM mining_c45");
3: while($rowGainRatio = mysql_fetch_array($sqlGainRatio)) {
4: $idGainRatio = $rowGainRatio['id'];
5: if ($rowGainRatio['inf_gain'] == 0 AND
  $rowGainRatio['split_info'] == 0){
6: $getGainRatio = 0;
7: } else {
8: $getGainRatio = round(($rowGainRatio['inf_gain'] /
  $rowGainRatio['split_info']),4);
9: mysql_query("UPDATE mining_c45 SET gain_ratio = $getGainRatio
  WHERE id = '$idGainRatio'");}
```

Segmen Program 5.3 merupakan potongan program untuk proses perhitungan gain ratio. Pada perhitungan ini terdapat looping untuk perhitungan gain ratio setiap atribut dan proses update nilai gain ratio pada database. Sub function insertAtributPohonKeputusan adalah fungsi yang menambahkan atribut dan information gain max ke DB pohon keputusan.

Segmen Program 5.5 Function insertAtributPohonKeputusan

```
1: function insertAtributPohonKeputusan($atribut, $nilai_atribut){
2:   $sqlInfGainMaxTemp = mysql_query("SELECT distinct atribut,
   gain_ratio FROM mining_c45 WHERE gain_ratio in (SELECT
   max(gain_ratio) FROM `mining_c45`) LIMIT 1");
3:   $rowInfGainMaxTemp = mysql_fetch_array($sqlInfGainMaxTemp);
4:   if ($rowInfGainMaxTemp['gain_ratio'] > 0) {
5:     $sqlInfGainMax = mysql_query("SELECT * FROM mining_c45 WHERE
   atribut = '$rowInfGainMaxTemp[atribut]'");
6:     while($rowInfGainMax = mysql_fetch_array($sqlInfGainMax)) {
7:       if ($rowInfGainMax['jml_kasus_tepat_waktu'] == 0 AND
   $rowInfGainMax['jml_kasus_tidak_tepat_waktu'] == 0) {
8:         $keputusan = 'Null';
9:       } else if ($rowInfGainMax['jml_kasus_tepat_waktu'] != 0 AND
   $rowInfGainMax['jml_kasus_tidak_tepat_waktu'] == 0) {
10:        $keputusan = 'Tepat Waktu'; } else if
   ($rowInfGainMax['jml_kasus_tepat_waktu'] == 0 AND
   $rowInfGainMax['jml_kasus_tidak_tepat_waktu'] != 0) {
11:        $keputusan = 'Tidak Tepat Waktu'; } else {
12:        $keputusan = '?'; }
13:     if (empty($atribut) AND empty($nilai_atribut)) {
14:       $kondisiAtribut = "AND $rowInfGainMax[atribut] =
   ~$rowInfGainMax[nilai_atribut]~";
15:       mysql_query("INSERT INTO pohon_keputusan_c45 VALUES ('',
   '$rowInfGainMax[atribut]', '$rowInfGainMax[nilai_atribut]', 0,
   '$rowInfGainMax[jml_kasus_tidak_tepat_waktu]',
   '$rowInfGainMax[jml_kasus_tepat_waktu]', '$keputusan', 'Belum',
   '$kondisiAtribut', 'Belum')");}
16:     else if (!empty($atribut) AND !empty($nilai_atribut)) {
17:       $sqlIdParent = mysql_query("SELECT id, atribut,
   nilai_atribut, jml_kasus_tepat_waktu,
   jml_kasus_tidak_tepat_waktu FROM pohon_keputusan_c45 WHERE
   atribut = '$atribut' AND nilai_atribut = '$nilai_atribut' order
   by id DESC LIMIT 1");
18:       $rowIdParent = mysql_fetch_array($sqlIdParent);
19:       mysql_query("INSERT INTO pohon_keputusan_c45 VALUES ('',
   '$rowInfGainMax[atribut]', '$rowInfGainMax[nilai_atribut]',
   $rowIdParent[id],
   '$rowInfGainMax[jml_kasus_tidak_tepat_waktu]',
   '$rowInfGainMax[jml_kasus_tepat_waktu]', '$keputusan', 'Belum',
   '', 'Belum')");}
20:   }}}loopingKondisiAtribut();}
```

Segmen program 5.5 adalah Fungsi untuk menginsertkan atribut dan information gain max ke DB pohon keputusan. Pertama-tama fungsi mengecek untuk inf gain tertinggi. Kemudian mengambil nilai atribut yang memiliki nilai inf gain tertinggi. Langkah selanjutnya terjadi pengecekan ,jika atribut yang diinput kosong (atribut awal) maka insert ke pohon keputusan id_parent = 0 dan Jika atribut yang diinput Tidak Tepat Waktu kosong maka insert ke pohon keputusan

dimana id_parent diambil dari tabel pohon keputusan sebelumnya (where atribut = atribut yang diinput).

Segmen Program 5.6 Function getInfGainMax

```
1: $sqlInfGainMaxAtribut = mysql_query("SELECT distinct atribut
FROM mining_c45 WHERE gain_ratio in (SELECT max(gain_ratio)
FROM `mining_c45`) LIMIT 1");
2: while($rowInfGainMaxAtribut =
mysql_fetch_array($sqlInfGainMaxAtribut)) {
3: $inf_gain_max_atribut = "$rowInfGainMaxAtribut[atribut]";
4: if (empty($atribut) AND empty($nilai_atribut)) {
5: loopingMiningC45($inf_gain_max_atribut);
6: } else if (!empty($atribut) AND !empty($nilai_atribut)) {
7: mysql_query("UPDATE pohon_keputusan_c45 SET diproses =
'Sudah' WHERE atribut = '$atribut' AND nilai_atribut =
'$nilai_atribut'");
8: loopingMiningC45($inf_gain_max_atribut); }}
```

Segmen program 5.6 adalah fungsi untuk memperoleh gain max untuk diproses dalam proses mining. Selanjutnya, akan dijelaskan mengenai function prepruning yang dipanggil pada function insertAtributPohonKeputusan.

Segmen Program 5.7 Function loopingPerhitunganPrePruning

```
1: function loopingPerhitunganPrePruning($positif, $negatif)
2: { $confidenceLevel = 95;
3: $z = (1 / 0.5775) * ($confidenceLevel / 100);
4: $z = round($z, 4);
5: $n = $positif + $negatif;
6: $n = round($n, 4);
7: if ($positif < $negatif) {
8: $r = $positif / ($n);
9: $r = round($r, 4);
10: return perhitunganPrePruning($r, $z, $n);
11: } elseif ($positif > $negatif) {
12: $r = $negatif / ($n);
13: $r = round($r, 4);
14: return perhitunganPrePruning($r, $z, $n);
15: } elseif ($positif == $negatif) {
16: $r = $negatif / ($n);
17: $r = round($r, 4);
18: return perhitunganPrePruning($r, $z, $n);
19: }
20: }
```

Segmen 5.7 adalah segmen program tentang perhitungan pre pruning, pada proses perhitungan prepruning ada variabel confidence level merupakan batas kepercayaan (confidence threshold).

Segmen Program 5.8 Function Perhitungan Kinerja

```

1: function countKeputusanKinerja($atribut1, $nilai_atribut1,
   $atribut2, $nilai_atribut2)
2:   $TepatWaktu = countKeputusan($atribut = 'keputusan_asli',
   $nilai_atribut = 'Tepat Waktu');
3:   $TidakTepatWaktu = countKeputusan($atribut =
   'keputusan_asli', $nilai_atribut = 'Tidak Tepat Waktu');
4:   $TepatWaktuC45IdentifikasiTepatWaktu =
   countKeputusanKinerja($atribut1 = 'keputusan_asli',
   $nilai_atribut1 = 'Tepat Waktu', $atribut2 = 'keputusan_c45',
   $nilai_atribut2 = 'Tepat Waktu');
5:   $TepatWaktuC45IdentifikasiTidakTepatWaktu =
   countKeputusanKinerja($atribut1 = 'keputusan_asli',
   $nilai_atribut1 = 'Tepat Waktu', $atribut2 = 'keputusan_c45',
   $nilai_atribut2 = 'Tidak Tepat Waktu');
6:   $TepatWaktuC45IdentifikasiNull =
   countKeputusanKinerja($atribut1 = 'keputusan_asli',
   $nilai_atribut1 = 'Tepat Waktu', $atribut2 = 'keputusan_c45',
   $nilai_atribut2 = 'Null');
7:   $TidakTepatWaktuC45IdentifikasiTepatWaktu =
   countKeputusanKinerja($atribut1 = 'keputusan_asli',
   $nilai_atribut1 = 'Tidak Tepat Waktu', $atribut2 =
   'keputusan_c45', $nilai_atribut2 = 'Tepat Waktu');
8:   $TidakTepatWaktuC45IdentifikasiTidakTepatWaktu =
   countKeputusanKinerja($atribut1 = 'keputusan_asli',
   $nilai_atribut1 = 'Tidak Tepat Waktu', $atribut2 =
   'keputusan_c45', $nilai_atribut2 = 'Tidak Tepat Waktu');
9:   $TidakTepatWaktuC45IdentifikasiNull =
   countKeputusanKinerja($atribut1 = 'keputusan_asli',
   $nilai_atribut1 = 'Tidak Tepat Waktu', $atribut2 =
   'keputusan_c45', $nilai_atribut2 = 'Null');
10:  if ($TepatWaktu != ($TepatWaktuC45IdentifikasiTepatWaktu +
   $TepatWaktuC45IdentifikasiTidakTepatWaktu)) {
11:    $TepatWaktuC45IdentifikasiTidakTepatWaktu +=
   $TepatWaktuC45IdentifikasiNull;
12:    if ($TidakTepatWaktu !=
   ($TidakTepatWaktuC45IdentifikasiTepatWaktu
   + $TidakTepatWaktuC45IdentifikasiTidakTepatWaktu)) {
13:      $TidakTepatWaktuC45IdentifikasiTepatWaktu +=
   $TidakTepatWaktuC45IdentifikasiNull;
14:    }
15:    $precision =
   round((($TidakTepatWaktuC45IdentifikasiTidakTepatWaktu
   / ($TepatWaktuC45IdentifikasiTidakTepatWaktu
   + $TidakTepatWaktuC45IdentifikasiTidakTepatWaktu) * 100), 2);
16:    $recall =
   round((($TidakTepatWaktuC45IdentifikasiTidakTepatWaktu
   / ($TidakTepatWaktuC45IdentifikasiTepatWaktu
   + $TidakTepatWaktuC45IdentifikasiTidakTepatWaktu) * 100), 2);
17:    $accuracy =
   round(((($TidakTepatWaktuC45IdentifikasiTidakTepatWaktu
   + $TepatWaktuC45IdentifikasiTepatWaktu) / ($TidakTepatWaktu
   + $TepatWaktu) * 100), 2);
18:  }

```


Segmen Program 5.8 adalah potongan program menilai kinerja dari program yang dibuat termasuk menghitung nilai recall, precision, dan accuracy dari program yang telah dibuat.

Segmen Program 5.9 Function Penentuan Keputusan

```

1: while($rowDataPenentuKeputusan =
mysql_fetch_array($sqlDataPenentuKeputusan)) {
2: $sqlRulePenentuKeputusan = mysql_query("SELECT id, atribut,
nilai_atribut FROM rule_penentu_keputusan);
3: while($rowRulePenentuKeputusan =
mysql_fetch_array($sqlRulePenentuKeputusan)) {
4: if (!empty($rowRulePenentuKeputusan)) {
5: if ($rowRulePenentuKeputusan['atribut'] ==
$rowDataPenentuKeputusan['atribut']) {
6: if ($rowRulePenentuKeputusan['nilai_atribut'] ==
$rowDataPenentuKeputusan['nilai_atribut']) {
7: $arrayPenentuKeputusanTemp['id'] =
$rowRulePenentuKeputusan['id'];
8: $arrayPenentuKeputusanTemp['cocok'] = "Ya";
9: $arrayPenentuKeputusan[] = $arrayPenentuKeputusanTemp;
10: } elseif ($rowRulePenentuKeputusan['nilai_atribut'] !=
$rowDataPenentuKeputusan['nilai_atribut']) {
11: $arrayPenentuKeputusanTemp['id'] =
$rowRulePenentuKeputusan['id'];
12: $arrayPenentuKeputusanTemp['cocok'] = "Tidak";
13: $arrayPenentuKeputusan[] = $arrayPenentuKeputusanTemp;
14: }}}}
15: foreach ($arrayPenentuKeputusan as
$arrayPenentuKeputusanUpdate) {
16: mysql_query("UPDATE rule_penentu_keputusan SET cocok =
'$arrayPenentuKeputusanUpdate[cocok]' where id =
$arrayPenentuKeputusanUpdate[id]");}

```

Segmen Program 5.9 adalah potongan program untuk penentuan keputusan terhadap data baru yang diinputkan.

5.1.2 Implementasi Desain Interface

Pada sub bab ini akan dibahas mengenai desain interface dari program yang dibuat. Desain Interface terdiri dari halaman login dan Halaman Website Utama. Halaman Login adalah halaman untuk user login agar bisa mengakses Website Utama. Gambar 5.1 adalah gambar halaman login. Pada halaman login terdapat textbox username dan textbox password. Ada 2 level user yaitu admin dan dosen. User level admin bisa mengakses proses mining hingga pembentukan

pohon keputusan, sedangkan dosen hanya bisa mengakses menu penentuan prediksi mahasiswa lulus tepat waktu atau tidak.



**SPK Prediksi Kelulusan
Menggunakan Pohon Keputusan C4.5**

Login Administrator

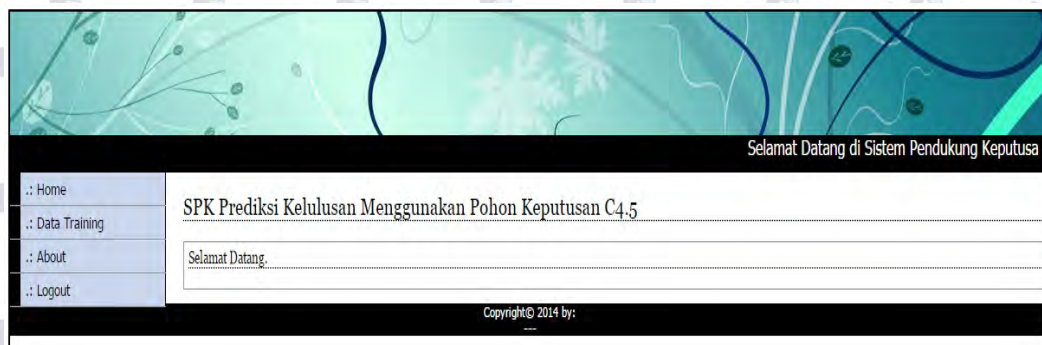
Username

Password

Contoh: admin

Gambar 5.1
Halaman Login

Halaman Website Utama memiliki beberapa pilihan menu di sebelah kiri. Gambar 5.2 adalah Halaman Website Utama.



Gambar 5.2
Halaman Website Utama

Halaman Website Utama Memiliki menu yang terdiri dari 4 menu pada awalnya, ketika user admin belum melakukan proses training data. Gambar 5.3 adalah gambar sub menu data Training yaitu sub menu untuk menginputkan data training ke dalam sistem.

Selamat Datang di Sistem Pendukung Keputusan - Menggunakan Poh

- Home
- Data Training
- Partisi Data
- Export Data
- C45 » Proses Mining
- About
- Logout

Data Training

Input Data Training

Data Ke:

Lama Studi:

Waktu status jurusan:

Nama PT S1:

Jurusan S1:

IPK S1:

IPK S2:

GMAT:

TOEFL:

WBJD:

Wawancara:

Skor Akhir:

Tahun masuk:

Tahun masuk s2:

Perak S1 ke S2:

Nama Perusahaan:

Lulus Sarjana Teltun:

Akreditasi s1:

program studi:

status program studi:

bidang keahlian:

biaya:

Class:

Status Data:

[« Prev](#)
[1](#)
[2](#)
[3](#)
[4](#)
[5](#)
[» Next](#)

NO	DATA KE	LAMA STUDI	Waktu status jurusan	NAMA PT S1	JURUSAN S1	IPK S1	IPK S2	GMAT	TOEFL	WBJD	WAWANCARA	SAKHIR	TAHUN MASUK	TAHUN MASUK S2	NAMA PERUSAHAAN	LULUS SARJANA Teltun	AKREDITASI S1	PROGRAM STUDI	STATUS PROGRAM STUDI
1	1	2	Membuka	Universitas terbuka	Statistika	3,6 B	2,91 C	54,94 D	213 B	72 B	80 C	69,88 D	qtr 2009/2010	2009	STIKOM Surabaya	2001	A	HE	Sebelang
2	2	3	MManajemen	Universitas Widyadarmas	Manajemen	3,28 B	3,42 B	66,93 C	130 D	80 B	80 C	70,28 D	qtr 2009/2010	2009	BNJ Jemberan	2003	B	MI	Sebelang

Gambar 5.3
Sub Menu Data Training

List Menu ditunjukkan pada gambar 5.3 nomor 1. List menu ini bertambah ketika user selesai menginputkan data Training ke sistem. Input data Training bisa dilakukan dengan 2 cara yaitu input satu per satu data gambar 5.3 nomor 2 atau dengan cara bulk import yaitu mengimport file cvs gambar 5.3. nomor 4. Gambar 5.3 nomor 3 adalah fungsi untuk menghapus semua data training. Gambar 5.3 nomor 5 adalah display data yang telah terimport ke dalam sistem dalam bentuk table dengan *pagination*. Ada fasilitas mengedit atau menghapus pada satu data tertentu. Menginputkan data training adalah dengan cara menekan button input pada input single data atau dengan cara menekan tombol browse memilih file yang akan diinputkan ke sistem lalu tekan button submit. Setelah data masuk ke database secara otomatis menu akan bertambah yaitu partisi data, ekspor data dan proses mining.

Selamat Datang di Sistem Pendukung Keputusan - Menggunakan

1 Data Training » Partisi Data

Set Data Training (Semua Data) : % **2** Proses

JUMLAH DATA (0 DATA)	DATA TRAINING (0 DATA)	DATA TESTING (0 DATA)
Tidak Tepat Waktu (0 Data)	0	0
Tepat Waktu (0 Data)	0	0

3

Copyright© 2014 by:

Gambar 5.4
Sub Menu Partisi Data

Setelah proses Input data Training selanjutnya adalah Partisi data. Pada sub menu ini, user dapat menginputkan persentase dari data yang telah terinput ke sistem untuk diperlakukan sebagai data training dan sisanya menjadi data testing. Gambar 5.4 no mor 2 adalah tempat menginputkan persentase dan menekan tombol proses untuk melakukan proses. Pemilihan data yang mana menjadi data training dan mana yang akan menjadi data testing dilakukan secara random. Pada gambar 5.4 terdapat sub menu export data, yaitu proses mengeksport data training dan testing.

C45 » Mining C4.5

Proses Mining C4.5 : Non Pruning ▼

Proses!

Gambar 5.5
Sub Menu Proses Mining C4.5

Gambar 5.5 merupakan gambar sub menu mining C45. Pada halaman ini, user dapat melakukan proses mining dengan pilihan menggunakan pre pruning, atau tanpa menggunakan pruning. Untuk melihat perhitungan gain dan entrophy dari proses mining yang telah dilakukan, user bisa memilih sub menu C45-Perhitungan lihat gambar 5.6 dan user dapat melihat pohon keputusan hasil dari proses mining dengan memilih sub menu C45 pohon keputusan lihat gambar 5.7.

C45 » Perhitungan C4.5

Opsi: [Hapus Semua Data](#)

NO	ATRIBUT GAIN RATIO MAX	ATRIBUT	NILAI ATRIBUT	JUMLAH KASUS TOTAL	JUMLAH KASUS TIDAK TEPAT WAKTU	JUMLAH KASUS TEPAT WAKTU	ENTROPY	INFORMATION GAIN	SPLIT INFO	GAIN RATIO
1	ips_s2_huruf	Total	Total	134	51	83	0.9585			0
2	ips_s2_huruf	akreditasi_s1	A	82	29	53	0.9373	0.0243	1.3091	0.0186
3	ips_s2_huruf	akreditasi_s1	B	36	12	24	0.9183	0.0243	1.3091	0.0186
4	ips_s2_huruf	akreditasi_s1	C	16	10	6	0.9544	0.0243	1.3091	0.0186
5	ips_s2_huruf	program_studi	MBM	2	2	0	0	0.042	1.635	0.0257
6	ips_s2_huruf	program_studi	MI	51	16	35	0.8974	0.042	1.635	0.0257
7	ips_s2_huruf	program_studi	MP	30	14	16	0.9968	0.042	1.635	0.0257
8	ips_s2_huruf	program_studi	MTK	1	1	0	0	0.042	1.635	0.0257
9	ips_s2_huruf	program_studi	MTI	50	18	32	0.9427	0.042	1.635	0.0257
10	ips_s2_huruf	status_program_studi	Sebidang	87	37	50	0.9838	0.0116	0.9347	0.0124
11	ips_s2_huruf	status_program_studi	tidak sebidang	47	14	33	0.8787	0.0116	0.9347	0.0124
12	ips_s2_huruf	bidang_keahlian	Eksekutif	30	14	16	0.9968	0.0065	0.7672	0.0085
13	ips_s2_huruf	bidang_keahlian	Profesional	104	37	67	0.9391	0.0065	0.7672	0.0085
14	ips_s2_huruf	biaya	Biaya Sendiri	108	43	65	0.9699	0.0168	0.8099	0.0207

Gambar 5.6
Sub Menu Perhitungan C45

Gambar 5.6 terdapat tabel dari perhitungan gain dan entropy dari setiap atribut dari setiap iterasi. Selain itu juga ada gain max dan jumlah total kasus tiap atribut. Gambar 5.7 adalah gambar pohon keputusan dan rule yang dibentuk dari proses mining. Sub Menu proses kinerja untuk memproses perhitungan kinerja dari hasil data mining yang di peroleh. Sedangkan sub menu kinerja perbandingan adalah halaman yang menampilkan perbandingan keputusan asli dengan keputusan yang dibuat oleh algoritma C45 lihat gambar 5.7.

Kinerja » Tabel Perbandingan Keputusan

Opsi: [Hapus Semua Data](#)

< Prev 1 | 2 | 3 | 4 | 5 | 3 | Next >

NO	NRP	NAMA PT S1	JURUSAN S1	IPK S1	IPS S1	IPS S2	GMAT	TOEFL	M.BID	WAWANCARA	S.AKHIR	JARAK S1-S2	AKREDITASI	PROGRAM STUDI	STATUS PROGRAM STUDI	BIDANG KEAHLIAN	BIAYA	KEPUTUSAN ASLI	KEPUTUSAN C4.5 ID RULE C4.5
1	9109201301	Universitas terbuka	Statistika	3.6 B	3.45 B	3.59 B	54.94 D	553 B	72 B	80 C	69.88 D	Jauh	B	MI	tidak sebidang	Profesional	Biaya Sendiri	Tepat Waktu	Tepat Waktu 56
2	9109201508	ITB	Teknik Kimia	3.64 B	3.45 B	3.54 B	98.9 A	480 C	84 B	81 C	83.27 C	Jauh	A	MI	tidak sebidang	Eksekutif	Biaya Sendiri	Tepat Waktu	Tepat Waktu 56
3	9109205327	Univ. Petra Surabaya	Teknik Informatika	3.71 B	3.6 B	3.75 A	80.91 B	550 B	60 C	76 C	73.07 D	Dekat	A	MTI	Sebidang	Profesional	Biaya Sendiri	Tepat Waktu	Tepat Waktu 19
4	9109205503	ITS Surabaya	Teknologi Informasi	3.42 C	3.3 C	3.67 A	58 C	420 D	58 C	78 C	60.4 E	Dekat	A	MTI	Sebidang	Eksekutif	Biaya Sendiri	Tepat Waktu	Tepat Waktu 35
5	9109205318	STAN	Akuntansi	3.58 B	3.4 B	3.5 B	85.91 B	487 C	46 D	90 B	69.97 D	Jauh	B	MTI	tidak sebidang	Profesional	Biaya Sendiri	Tepat Waktu	Tidak Tepat Waktu 55

Gambar 5.7
Sub Menu Kinerja Perbandingan

Gambar 5.8 a adalah gambar halaman kinerja table penilaian. Pada halaman ini ditampilkan hasil perhitungan nilai recall, hasil perhitungan nilai precision, dan hasil perhitungan accuracy.

Kinerja » Tabel Penilaian

Tabel Penilaian C4.5

	Diidentifikasi Tidak Tepat Waktu oleh C4.5	Diidentifikasi Tepat Waktu oleh C4.5
Keputusan Asli: Tidak Tepat Waktu = 15	6	9
Keputusan Asli: Tepat Waktu = 29	3	26

1. Precision = $6 / (3 + 6) * 100\% = 66.67 \%$
2. Recall = $6 / (9 + 6) * 100\% = 40 \%$
3. Accuracy = $(6 + 26) / (15 + 29) * 100\% = 72.73 \%$

Gambar 5.8
Sub Menu Kinerja Tabel Penilaian

Gambar 5.9 a adalah gambar halaman untuk Penentuan Keputusan. Pada halaman ini terdapat 4 bagian input penentuan keputusan single data pada gambar 5.9 nomor 1, input multiple data pada gambar 5.9 nomor 2, untuk hapus data semuanya di table penentuan data pada gambar 5.9 nomor 3. Tabel Penentuan Keputusan pada gambar 5.9 nomor 4.

Penentu Keputusan

Input Data

NRP

Nama PT SL

Jurusan SL

IPK SL

IPS SL

IPS SZ

Gmat

Toefl

M.BID

Wawancara

Skor Akhir

Jarak SL-SZ

Program Studi

Bidang Keahlian

Biaya

Input

Amatir File .csv

Pilih File

Tidak ada file yang dipilih

Submit

Ops: Hapus Semua Data

NO	NRP	NAMA PT SL	JURUSAN SL	IPK SL	IPS SL	IPS SZ	GMAT	TOEFL	M.BID	WAWANCARA	SAKHIR	RAK SL-SZ	AKREDITASI	PROGRAM STUDI	STATUS PROGRAM STUDI	BIDANG KEAHLIAN	BIAYA	KEPUTUSAN C4.5 ID RULE	OPSI
1	9109201301	Universitas terbuka	Statistika	3,6 B	3,15 B	3,59 B	54,91 D	553 B	72 B	80 C	69,88 D	Jauh	B	MI	tidak sebanding	Profesional	Biaya Sendiri	Tepat Waktu 56	Hapus
2	9109201307	Universitas Widyadarmas	Manajemen	3,28 D	3,09 C	3,15 C	66,93 C	430 D	80 B	80 C	70,28 D	Jauh	B	MI	Sebanding	Eksekutif	Biaya Sendiri	Tepat Waktu 122	Hapus
3	9109201310	ITS Surabaya	Teknik Kelautan	3,46 C	3,27 C	3,14 C	82,93 B	477 C	80 B	77 C	76,27 C	Jauh	A	MI	tidak sebanding	Eksekutif	Biaya Sendiri	Tepat Waktu 125	Hapus
4	9109201308	ITB	Teknik Kimia	3,64	3,15	3,51	98,9	480	81	81	83,27	Jauh	A	MI	tidak sebanding	Eksekutif	Biaya Sendiri	Tepat Waktu 125	Hapus

Gambar 5.9
Sub Menu Penentu Keputusan

5.2 Uji Coba Sistem

Pada Sub Bab ini, akan dijelaskan mengenai hasil ujicoba yang dilakukan terhadap sistem. Uji coba di dilakukan dari contoh data yang diperoleh dengan beberapa skenario dan beberapa variasi partisi data. Penelitian ini di ujicoba dengan 3 macam skenario yaitu skenario pendeteksian saat mahasiswa baru diterima sebagai mahasiswa MMT ITS, pendeteksian saat mahasiswa telah menyelesaikan semester 1 di MMT ITS, dan pendeteksian saat mahasiswa telah menyelesaikan semester 2 di MMT ITS.

Sedangkan untuk partisi data, penelitian Skenario I, Skenario II, dan Skenario II di ujicoba dengan 3 variasi partisi data yaitu 70: 30, 80: 20 dan 90 : 10. Partisi data tersebut dibedakan berdasarkan perbandingan persentase data training dan data testing yang diujikan. Partisi data 70 : 30 artinya 70% dari contoh data diperlakukan sebagai data training dan 30% dari contoh data diperlakukan sebagai data testing. Partisi data 80 : 20 artinya 80% dari contoh data diperlakukan sebagai data training dan 20% dari contoh data diperlakukan sebagai data testing. Partisi data 90 : 10 artinya 90% dari contoh data diperlakukan sebagai data training dan 10% dari contoh data diperlakukan sebagai data testing. Pemilihan data yang akan jadi kandidat training maupun testing adalah bersifat Random.

Selain itu juga dilakukan uji coba dengan mengelompokan data berdasarkan program studi. Program studi yang akan di ujicoba adalah program studi MTI, MP dan MI. sedangkan untuk 2 program studi lainnya tidak di ujicoba karena minimnya data mahasiswa pada program studi tersebut. Pada uji coba program studi ini , partisi data yang digunakan adalah partisi data 90:10 dengan perulangan ujicoba sebanyak 5 kali.

5.2.1 Uji Coba Skenario I

Uji coba skenario I mengilustrasikan pada saat mahasiswa, baru diterima menjadi mahasiswa MMT ITS. Pengukuran hasil uji coba diukur dengan nilai precision, nilai recall dan nilai accuracy. Berikut akan disajikan dalam tabel nilai precision, recall dan accurasi dari hasil ujicoba yang telah dilakukan :

Tabel 5.1
Hasil Precision Uji Coba Skenario I

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	33.33%	37.00%	60.00%	28.57%	25.00%	66.67%	41.76%
II	44.83%	40.00%	50.00%	55.56%	50.00%	40.00%	46.73%
III	33.33%	28.57%	66.67%	75.00%	50.00%	100.00%	58.93%
IV	42.86%	30.00%	40.00%	50.00%	26.67%	50.00%	39.92%
V	27.27%	26.67%	50.00%	40.00%	20.00%	100.00%	43.99%
AVG	36.32%	32.45%	53.33%	49.83%	34.33%	71.33%	46.27%

Tabel 5.1 adalah Tabel Hasil Precision Uji Coba Skenario I. Pada skenario I ini dapat dilihat pada proses klasifikasi tanpa pruning, terjadi peningkatan persentase data training diikuti kecenderungan peningkatan nilai precision. Hal ini berarti peningkatan jumlah data training berpengaruh positif untuk peningkatan nilai precision. Sedangkan untuk klasifikasi menggunakan pre pruning, peningkatan jumlah data training tidak selalu meningkat nilai precision, terkadang malah membuat nilai precision menjadi lebih kecil. Hal ini diperkirakan disebabkan oleh keterbatasan data yang dimiliki peneliti untuk proses ujicoba menyebabkan pembentukan rule keputusan yang kurang sempurna pada proses klasifikasi dengan pre pruning. Pada Skenario ini, memiliki kecenderungan nilai precision menjadi lebih besar ketika dilakukan proses pre pruning.

Tabel 5.2
Hasil Recall Uji Coba Skenario I

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	31.58%	27.27%	37.50%	21.05%	18.18%	25.00%	26.76%
II	61.90%	46.15%	50.00%	23.81%	15.38%	33.33%	38.43%
III	12.00%	33.30%	25.00%	12.00%	12.50%	12.50%	17.88%
IV	50.00%	75.00%	40.00%	11.11%	50.00%	20.00%	41.02%
V	37.50%	30.77%	60.00%	12.50%	7.69%	40.00%	31.41%
AVG	38.60%	42.50%	42.50%	16.09%	20.75%	26.17%	31.10%

Tabel 5.2 adalah Tabel hasil Recall Uji Coba Skenario I. dari tabel tersebut dapat disimpulkan bahwa proses pre-pruning menyebabkan nilai Recall menurun. Pada proses Non Pruning maupun proses Pre Pruning, dapat dilihat bahwa bertambahnya data bisa menyebabkan nilai recall cenderung meningkat. Hasil recall menurun ketika

Tabel 5.3
Hasil Accuracy Uji Coba Skenario I

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	52.83%	63.89%	61.11%	52.83%	58.33%	61.11%	58.35%
II	54.72%	55.56%	66.67%	62.26%	63.89%	61.11%	60.70%
III	47.17%	50.00%	61.11%	56.60%	66.67%	61.11%	57.11%
IV	60.38%	55.56%	66.67%	66.04%	58.33%	72.22%	63.20%
V	50.94%	44.44%	72.22%	67.92%	55.56%	83.30%	62.40%
AVG	53.21%	53.89%	65.56%	61.13%	60.56%	67.77%	60.35%

Tabel 5.3 adalah Tabel hasil Accuracy Uji Coba Skenario I. Pada scenario I ini, proses pre pruning bisa meningkatkan nilai Accuracy, terbukti bahwa hampir semua nilai Accuracy pada partisi data bersesuaian yang disajikan tabel 5.3 mengalami peningkatan ketika dilakukan pre pruning. Peningkatan jumlah data pada scenario I, memiliki kecenderungan diikuti oleh peningkatan nilai Accuracy. Hal ini dapat disimpulkan bahwa bertambahnya data training meningkatkan nilai Accuracy hasil uji coba.

5.2.2 Uji Coba Skenario II

Uji coba skenario II mengilustrasikan pada saat mahasiswa, telah menyelesaikan semester I di MMT ITS. Skenario ini terdapat 1 variabel mining tambahan yaitu Index Prestasi Semester 1. Pengukuran hasil uji coba diukur dengan nilai precision, nilai recall dan nilai accuracy. Berikut akan disajikan dalam tabel nilai precision, recall dan accuracy dari hasil ujicoba yang telah dilakukan :

Tabel 5.4
Hasil Precision Uji Coba Skenario II

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	36.84%	30.00%	71.43%	33.33%	25.00%	66.67%	43.88%
II	46.15%	41.67%	50.00%	55.56%	50.00%	40.00%	47.23%
III	40.00%	36.36%	66.67%	75.00%	50.00%	100.00%	61.34%
IV	44.00%	33.33%	33.30%	50.00%	26.67%	50.00%	39.55%
V	27.27%	28.57%	50.00%	40.00%	20.00%	100.00%	44.31%
AVG	38.85%	33.99%	54.28%	50.78%	34.33%	71.33%	47.26%

Tabel 5.4 adalah Tabel Hasil Precision Uji Coba Skenario II. Pada skenario II ini dapat dilihat pada proses klasifikasi tanpa pruning, terjadi peningkatan persentase data training diikuti kecenderungan peningkatan nilai precision. Sedangkan untuk klasifikasi menggunakan pre pruning, peningkatan jumlah data training tidak selalu meningkat nilai precision, terkadang malah membuat nilai precision menjadi lebih kecil.

Tabel 5.5
Hasil Recall Uji Coba Skenario II

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	36.84%	27.27%	62.50%	26.32%	18.18%	25.00%	32.69%
II	57.14%	38.46%	50.00%	23.81%	15.38%	33.33%	36.35%
III	16.00%	33.33%	25.00%	12.00%	16.67%	12.50%	19.25%
IV	61.11%	75.00%	40.00%	11.11%	56.00%	20.00%	43.87%
V	37.50%	30.77%	40.00%	12.50%	7.69%	40.00%	28.08%
AVG	41.72%	40.97%	43.50%	17.15%	22.78%	26.17%	32.05%

Tabel 5.5 adalah Tabel hasil Recall Uji Coba Skenario II. Dari tabel tersebut dapat disimpulkan bahwa proses pre-pruning menyebabkan nilai Recall menurun. Pada proses Non Pruning maupun proses Pre Pruning, dapat dilihat bahwa bertambahnya data bisa menyebabkan nilai recall bertambah maupun berkurang. Hal ini bisa jadi disebabkan oleh variasi rule yang dibentuk dari proses klasifikasi data training yang dirandom.

Tabel 5.6
Hasil Accuracy Uji Coba Skenario II

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	54.72%	58.33%	72.22%	54.72%	58.33%	61.11%	59.91%
II	56.60%	58.33%	66.67%	62.26%	63.89%	61.11%	61.48%
III	49.06%	58.33%	61.10%	56.60%	66.67%	61.11%	58.81%
IV	60.38%	61.11%	61.11%	66.04%	58.33%	72.22%	63.20%
V	50.94%	47.22%	72.22%	67.92%	55.56%	83.33%	62.87%
AVG	54.34%	56.66%	66.66%	61.51%	60.56%	67.78%	61.25%

Tabel 5.6 adalah Tabel hasil Accuracy Uji Coba Skenario II. Pada scenario II ini, proses pre pruning bisa meningkatkan nilai Accuracy, terbukti bahwa hampir semua nilai Accuracy pada partisi data bersesuaian yang disajikan tabel 5.6 mengalami peningkatan ketika dilakukan pre pruning. Peningkatan jumlah data pada skenario II, memiliki kecenderungan diikuti oleh peningkatan nilai *Accuracy*. Hal ini dapat disimpulkan bahwa bertambahnya data training meningkatkan nilai Accuracy hasil uji coba.

5.2.3 Uji Coba Skenario III

Uji coba skenario III mengilustrasikan pada saat mahasiswa, telah menyelesaikan semester I dan semester II di MMT ITS. Skenario ini terdapat 2 variabel mining tambahan yaitu Index Prestasi Semester 1 dan Index Prestasi Semester 2. Skenario I ini di uji coba menggunakan 3 macam partisi data, masing-masing partisi dilakukan percobaan sebanyak 5 kali. Partisi data 70 : 30 artinya 70% dari contoh data yang ada diperlakukan sebagai data training dan 30% dari contoh data yang ada diperlakukan sebagai data testing, penentuan bersifat random. Partisi data 80 : 20 artinya 80% dari contoh data yang ada diperlakukan sebagai data training dan 20% dari contoh data yang ada diperlakukan sebagai data testing, penentuan bersifat random. Partisi data 90 : 10 artinya 90% dari contoh data yang ada diperlakukan sebagai data training dan 10% dari contoh data yang ada diperlakukan sebagai data testing, penentuan bersifat random. Pengukuran hasil uji coba diukur dengan nilai precision, nilai recall dan nilai

accuracy. Berikut akan disajikan dalam tabel nilai precision, recall dan accurasi dari hasil ujicoba yang telah dilakukan :

Tabel 5.7
Hasil Precision Uji Coba Skenario III

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	40.91%	9.09%	57.14%	33.33%	25.00%	57.14%	37.10%
II	48.15%	41.67%	42.86%	55.56%	50.00%	40.00%	46.37%
III	41.67%	45.95%	75.00%	75.00%	50.00%	100.00%	64.60%
IV	42.86%	31.58%	42.86%	50.00%	26.67%	50.00%	40.66%
V	31.82%	44.00%	60.00%	40.00%	20.00%	100.00%	49.30%
AVG	41.08%	34.46%	55.57%	50.78%	34.33%	69.43%	47.61%

Tabel 5.7 adalah Tabel Hasil Precision Uji Coba Skenario III. Pada skenario III ini dapat dilihat pada proses klasifikasi tanpa pruning, terjadi peningkatan persentase data training diikuti kecenderungan peningkatan nilai precision. Sedangkan untuk klasifikasi menggunakan pre pruning, peningkatan jumlah data training tidak selalu meningkat nilai precision, terkadang malah membuat nilai precision menjadi lebih kecil.

Tabel 5.8
Hasil Recall Uji Coba Skenario III

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	47.37%	9.09%	50.00%	26.32%	18.18%	50.00%	33.49%
II	61.90%	38.46%	50.00%	23.81%	15.38%	33.33%	37.15%
III	20.00%	41.67%	37.50%	12.00%	16.67%	12.50%	23.39%
IV	50.00%	75.00%	60.00%	11.11%	56.00%	20.00%	45.35%
V	43.75%	61.54%	60.00%	12.50%	7.69%	40.00%	37.58%
AVG	44.60%	45.15%	51.50%	17.15%	22.78%	31.17%	35.39%

Tabel 5.8 adalah Tabel hasil Recall Uji Coba Skenario III. Dari tabel tersebut dapat disimpulkan bahwa proses pre-pruning menyebabkan nilai Recall menurun. Pada proses Non Pruning maupun proses Pre Pruning, dapat dilihat

bahwa bertambahnya data bisa menyebabkan nilai recall bertambah maupun berkurang. Hal ini bisa jadi disebabkan oleh variasi rule yang dibentuk dari proses klasifikasi data training yang dirandom.

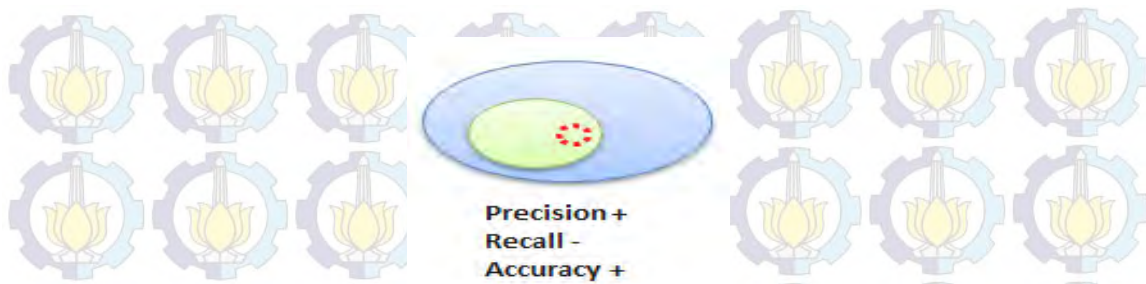
Tabel 5.9
Hasil Accuracy Uji Coba Skenario III

UJI COBA	NON-PRUNING			PRE - PRUNING			AVG
	70:30	80:20	90:10	70:30	80:20	90:10	
I	56.60%	44.44%	61.11%	54.72%	58.33%	61.11%	56.05%
II	58.49%	58.33%	61.11%	62.26%	63.89%	61.11%	60.87%
III	49.06%	63.89%	66.67%	56.60%	66.67%	61.11%	60.67%
IV	60.38%	58.33%	66.67%	66.04%	58.33%	72.22%	63.66%
V	54.72%	58.33%	77.78%	67.92%	55.56%	83.33%	66.27%
AVG	55.85%	56.66%	66.67%	61.51%	60.56%	67.78%	61.50%

Tabel 5.9 adalah Tabel hasil Accuracy Uji Coba Skenario III. Pada skenario III ini, proses pre pruning bisa meningkatkan nilai Accuracy, terbukti bahwa hampir semua nilai Accuracy pada partisi data bersesuaian yang disajikan tabel 5.9 mengalami peningkatan ketika dilakukan pre pruning. Peningkatan jumlah data pada skenario III, memiliki kecenderungan diikuti oleh peningkatan nilai Accuracy. Hal ini dapat disimpulkan bahwa bertambahnya data training meningkatkan nilai Accuracy hasil uji coba.

Dari data-data dan kesimpulan dari masing - masing tabel di atas dapat ditarik kesimpulan bahwa Pada skenario I,II,dan III :

1. Proses Pre Pruning menyebabkan nilai accuracy meningkat, nilai precision meningkat dan nilai recall menurun. Menurunnya nilai recal disebabkan oleh karena pada proses pre pruning, rule yang dibentuk lebih pendek karena proses pruning sehingga ruang pendektesian jadi lebih kecil digambarkan pada gambar 5.10. Lingkaran merah adalah hasil klarifikasi mahasiswa yang memiliki berpotensi lulus tidak tepat waktu.

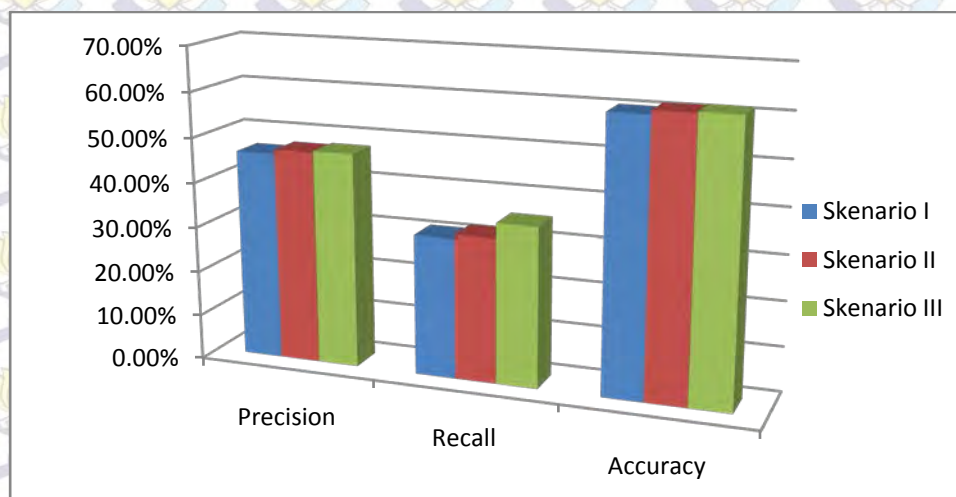


Gambar 5.10
Accuracy, Recall dan Precision

2. Bertambahnya data training menyebabkan nilai Accuracy, nilai Precision, dan nilai Recall memiliki kecenderungan meningkat.

5.2.4 Perbandingan Hasil Uji Coba Skenario I , Skenario II, Skenario III

Pada Sub bab ini, dibahas mengenai perbandingan hasil uji coba skenario I , skenario II, dan skenario II. Perbedaan tiga skenario tersebut adalah variable Index Prestasi Semester 1 (IPS1) dan Index Prestasi Semester 2(IPS 2). Pada skenario 1 IPS1 maupun IPS2 tidak ikut sebagai variable mining. Pada skenario II, variabel IPS1 ikut dalam proses mining karena mahasiswa telah menempuh studi di MMT ITS selama satu semester. Pada Skenario III, variabel IPS1 dan IPS2 ikut dalam proses mining karena mahasiswa telah menempuh studi di MMT ITS selama 2 semester yaitu pada semester 1 dan semester 2 sehingga mahasiswa telah memiliki kedua nilai tersebut.



Gambar 5.11
Perbandingan skenario I, skenario II, Skenario III

Pada gambar 5.11 dapat dilihat bahwa terjadi sedikit peningkatan Precision, Recall dan accuracy ketika variabel IPS1 dan IPS2 diikutkan dalam proses mining. Dugaan awal yang memperkirakan adanya peranan IPS1 dan IPS2 dalam pendeteksian lulus tidak tepat waktu adalah benar, walaupun pada percobaan ini peranan yang diberikan tergolong kecil. Hal ini membuktikan bahwa nilai IPS 1 dan nilai IPS2 mempunyai peranan yang cukup signifikan terhadap ketepatan lulus mahasiswa dalam menyelesaikan studinya di MMT ITS.

5.2.5 Uji Coba Program Studi MTI

Pada Sub bab ini dibahas mengenai hasil uji coba program studi MTI. Pada Uji Coba ini menggunakan data training mahasiswa yang mengambil program studi MTI. Banyak data yang digunakan dalam uji coba ini adalah sebanyak 67 data.

Tabel 5.10
Hasil Uji Coba Program Studi MTI

SKN	Uji Coba	NON-PRUNING			PRE - PRUNING		
		Precision	Recall	Acc.	Precision	Recall	Acc.
I	VI	40.00%	100.00%	57.14%	50.00%	100.00%	71.43%
	VII	33.33%	100.00%	71.43%	25.00%	100.00%	57.14%
	VIII	33.33%	50.00%	57.14%	100.00%	50.00%	85.71%
	IX	33.33%	33.33%	42.86%	33.33%	33.33%	42.86%
	X	100.00%	25.00%	57.14%	50.00%	25.00%	42.86%
II	VI	33.33%	100.00%	42.86%	50.00%	100.00%	71.43%
	VII	33.33%	100.00%	71.43%	25.00%	100.00%	57.14%
	VIII	33.33%	50.00%	57.14%	100.00%	50.00%	85.71%
	IX	33.33%	33.33%	42.86%	33.33%	33.33%	42.86%
	X	100.00%	25.00%	57.14%	50.00%	25.00%	42.86%
III	VI	40.00%	100.00%	57.14%	66.67%	100.00%	85.71%
	VII	33.33%	100.00%	71.43%	25.00%	100.00%	57.14%
	VIII	100.00%	50.00%	85.71%	100.00%	50.00%	85.71%
	IX	33.33%	33.33%	42.86%	33.33%	33.33%	42.86%
	X	100.00%	25.00%	57.14%	50.00%	25.00%	42.86%
AVG		52.00%	61.67%	58.09%	52.78%	61.67%	60.95%

Pada Tabel 5.10 dapat dilihat hasil uji coba menunjukan dengan dikelompokkannya berdasarkan program studi MTI maka nilai recall menjadi lebih besar dan cenderung mendekat sempurna, dan nilai Precisionnya menjadi lebih tinggi dibanding dengan percobaan sebelumnya yang melibatkan semua program studi. Ketika variabel nilai IPS1 dan IPS2 ikut dalam proses mining nilai akurasi cenderung meningkat, maka dapat disimpulkan bahwa pada mahasiswa program studi MTI nilai IPS 1 dan IPS 2 memberikan pengaruh yang cukup signifikan terhadap ketepatan lulus mahasiswa.

5.2.5 Uji Coba Program Studi MI

Pada Sub bab ini dibahas mengenai hasil uji coba program studi MI. Pada Uji Coba ini menggunakan data training mahasiswa yang mengambil program studi MTI saja. Dengan partisi data yang digunakan dalam uji coba ini adalah 90:10 dan dilakukan perulangan uji coba sebanyak 3 kali dengan 3 macam Skenario. Skenario sama dengan skenario pada bab bab sebelumnya. Banyak data yang digunakan dalam uji coba ini adalah sebanyak 73 data.

Tabel 5.11
Hasil Uji Coba Program Studi MI

SKN	Uji Coba	NON-PRUNING			PRE - PRUNING		
		Precision	Recall	Acc.	Precision	Recall	Acc.
I	XI	33.33%	100.00%	71.43%	50.00%	100.00%	85.71%
	XII	25.00%	50.00%	42.86%	50.00%	100.00%	85.71%
	XIII	33.33%	25.00%	28,57%	100.00%	25.00%	57.14%
	XIV	50.00%	66.67%	57.14%	100.00%	33.33%	71.43%
	XV	100.00%	50.00%	85.71%	100.00%	50.00%	85.71%
II	XI	33.33%	100.00%	71.43%	50.00%	100.00%	85.71%
	XII	33.33%	50.00%	57.14%	100.00%	50.00%	85.71%
	XIII	50.00%	25.00%	42.86%	100.00%	25.00%	57.14%
	XIV	50.00%	66.67%	57.14%	100.00%	33.33%	71.43%
	XV	100.00%	50.00%	85.71%	100.00%	50.00%	85.71%
III	XI	25.00%	100.00%	57.14%	50.00%	100.00%	85.71%
	XII	25.00%	50.00%	42.86%	100.00%	50.00%	85.71%
	XIII	50.00%	25.00%	42.86%	100.00%	25.00%	57.14%
	XIV	50.00%	66.67%	57.14%	100.00%	33.33%	71.43%

SKN	Uji Coba	NON-PRUNING			PRE - PRUNING		
		Precision	Recall	Acc.	Precision	Recall	Acc.
	XV	100.00%	50.00%	85.71%	100.00%	50.00%	85.71%
	AVG	50.55%	58.33%	61.22%	86.67%	55.00%	77.14%

Pada Tabel 5.11 dapat dilihat hasil uji coba menunjukkan dengan dikelompokkannya berdasarkan program studi MI maka nilai Precision menjadi lebih besar, dan nilai Recallnya menjadi lebih tinggi dibanding dengan percobaan sebelumnya yang melibatkan semua program studi. Sedangkan nilai Accuracy menjadi lebih baik. Ketika variabel nilai IPS1 dan IPS2 ikut dalam proses mining nilai akurasi menjadi turun, maka dapat disimpulkan bahwa pada mahasiswa program studi MTI nilai IPS 1 dan IPS 2 tidak memberikan pengaruh yang signifikan terhadap ketepatan lulus mahasiswa.

5.2.5 Uji Coba Program Studi MP

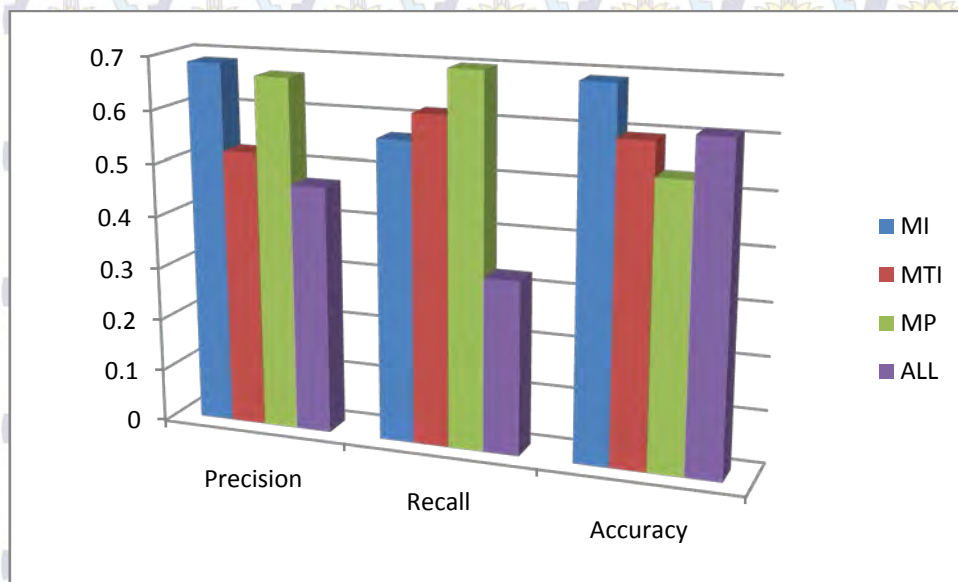
Pada Sub bab ini dibahas mengenai hasil uji coba program studi MP. Pada Uji Coba ini menggunakan data training mahasiswa yang mengambil program studi MTI saja. Dengan partisi data yang digunakan dalam uji coba ini adalah 90:10 dan dilakukan perulangan uji coba sebanyak 3 kali dengan 3 macam Skenario. Skenario sama dengan skenario pada bab sebelumnya. Banyak data yang digunakan dalam uji coba ini adalah sebanyak 33 data.

Tabel 5.12
Hasil Uji Coba Program Studi MP

SKN	Uji Coba	NON-PRUNING			PRE - PRUNING		
		Precision	Recall	Acc.	Precision	Recall	Acc.
I	XVI	100.00%	50.00%	66.67%	100.00%	50.00%	66.67%
	XVII	50.00%	50.00%	33.33%	50.00%	50.00%	33.33%
	XVIII	33.33%	100.00%	33.33%	33.33%	100.00%	33.33%
	XIX	100.00%	50.00%	66.67%	100.00%	50.00%	66.67%
	XX	50.00%	100.00%	66.67%	50.00%	100.00%	66.67%
II	XVI	100.00%	50.00%	66.67%	100.00%	50.00%	66.67%
	XVII	50.00%	50.00%	33.33%	50.00%	50.00%	33.33%
	XVIII	33.33%	100.00%	33.33%	33.33%	100.00%	33.33%
	XIX	100.00%	50.00%	66.67%	100.00%	50.00%	66.67%

SKN	Uji Coba	NON-PRUNING			PRE - PRUNING		
		Precision	Recall	Acc.	Precision	Recall	Acc.
	XX	50.00%	100.00%	66.67%	50.00%	100.00%	66.67%
III	XVI	100.00%	50.00%	66.67%	100.00%	50.00%	66.67%
	XVII	50.00%	50.00%	33.33%	50.00%	50.00%	33.33%
	XVIII	33.33%	100.00%	33.33%	33.33%	100.00%	33.33%
	XIX	100.00%	50.00%	66.67%	100.00%	50.00%	66.67%
	XX	50.00%	100.00%	66.67%	50.00%	100.00%	66.67%
	AVG	66.67%	70.00%	53.33%	66.67%	70.00%	53.33%

Pada Tabel 5.12 dapat dilihat hasil uji coba menunjukkan dengan dikelompokkannya berdasarkan program studi MP maka nilai Precision menjadi lebih besar, dan memiliki nilai recallnya meningkat. Sedangkan untuk nilai accuracy menjadi lebih rendah dari percobaan yang melibatkan semua program studi. Hal ini mungkin disebabkan data training untuk program studi MP kurang. Nilai Accuracy, Recall dan Precision tetap walaupun mengikuti variabel IPS1 dan IPS2 maupun diproses Pre Pruning.



Gambar 5.12
Perbandingan Program Studi MI, MTI, MP dan ALL

Pada gambar 5.12 nilai yang disajikan adalah nilai rata-rata, dapat dilihat bahwa perlakuan pengelompokan berdasarkan program studi dapat meningkatkan nilai Precision, Recall dan Accuracy.

BAB VI

KESIMPULAN DAN SARAN

Setelah proses desain, dan implementasi pada bab-bab sebelumnya maka pada bab ini dibahas mengenai kesimpulan dan saran yang berkaitan dengan pembuatan aplikasi pendeteksi kecenderungan lulus tidak tepat waktu pada Tesis ini. Kesimpulan didapat setelah mempertimbangkan beberapa kemudahan dan hambatan yang didapat selama proses pengerjaan. Sedangkan saran diberikan untuk pengembangan aplikasi pendeteksi kecenderungan lulus tidak tepat waktu lebih lanjut.

6.1 Kesimpulan

Pada sub bab ini dijelaskan mengenai kesimpulan-kesimpulan dari pembuatan aplikasi pendeteksi kecenderungan lulus tidak tepat waktu untuk penerapan algoritma klasifikasi C4.5 yang dibuat dengan menggunakan PHP. Kesimpulan ini didapatkan berdasarkan perkembangan selama proses pembuatan aplikasi. Kesimpulan-kesimpulan tersebut antara lain:

- Algoritma C4.5 dapat digunakan untuk memudahkan dalam pengambilan keputusan dengan memproyeksikan data-data yang ada ke dalam bentuk pohon keputusan, berdasarkan nilai entropy dan gain yang dimiliki masing-masing atribut data.
- Untuk hasil prediksi yang lebih akurat dibutuhkan data dalam jumlah besar, artinya semakin besar jumlah data yang digunakan maka semakin akurat hasil prediksi yang dihasilkan.
- Dari hasil ujicoba yang telah dilakukan, klasifikasi dengan prepruning confidence level 95%, membuat rule semakin pendek dan terkadang membuat recall, precision dan accuracy menjadi lebih jelek. Hal ini disebabkan karena kurangnya data training.
- Dengan jumlah data training dan jumlah data testing yang sama, namun beda variasi data dapat mengakibatkan perbedaan recall, precision dan accuracy, serta rule yang dibentuknya. Oleh karena itu, kualitas data training yang di

inputkan untuk klasifikasi memberikan kontribusi yang besar dalam pembentukan rule pada algoritma C4.5, selain jumlah data.

- Dari percobaan global seluruh program studi yang dilakukan, nilai IPS1 dan IPS2 dapat meningkatkan nilai precision, recall dan accuracy .
- Nilai accuracy, precision, dan recall cenderung meningkat ketika dilakukan permissahan per program studi di klasifikasi sendiri.

6.2 Saran

Berikut adalah beberapa saran yang dapat diberikan. Saran ini diambil berdasarkan perkembangan proses pendeteksian kecenderungan lulus tidak tepat waktu dengan algoritma C4.5 sebagai algoritma klasifikasinya. Saran-saran tersebut antara lain:

- Untuk pengembangan sistem sebaiknya membuat tampilan yang lebih baik dan lebih menarik dari sistem yang sudah ada karena dirasa tampilan sistem yang ada sekarang ini sangat sederhana.
- Sebaiknya data yang ditambahkan ke dalam sistem di update secara berkala, sehingga keputusan yang dihasilkan akurat.
- Untuk pendeteksian kecenderungan lulus tidak tepat waktu untuk mahasiswa MMT ITS, diperlukan tambahan variabel tanggal proposal tesis dan tanggal maju sidang tesis.

DAFTAR PUSTAKA

Witten, I. H., Frank, E., Hall, M. A. (2011). *Data Mining Practical Machine Learning Tools and Techniques* (3rd ed). USA: Elsevier

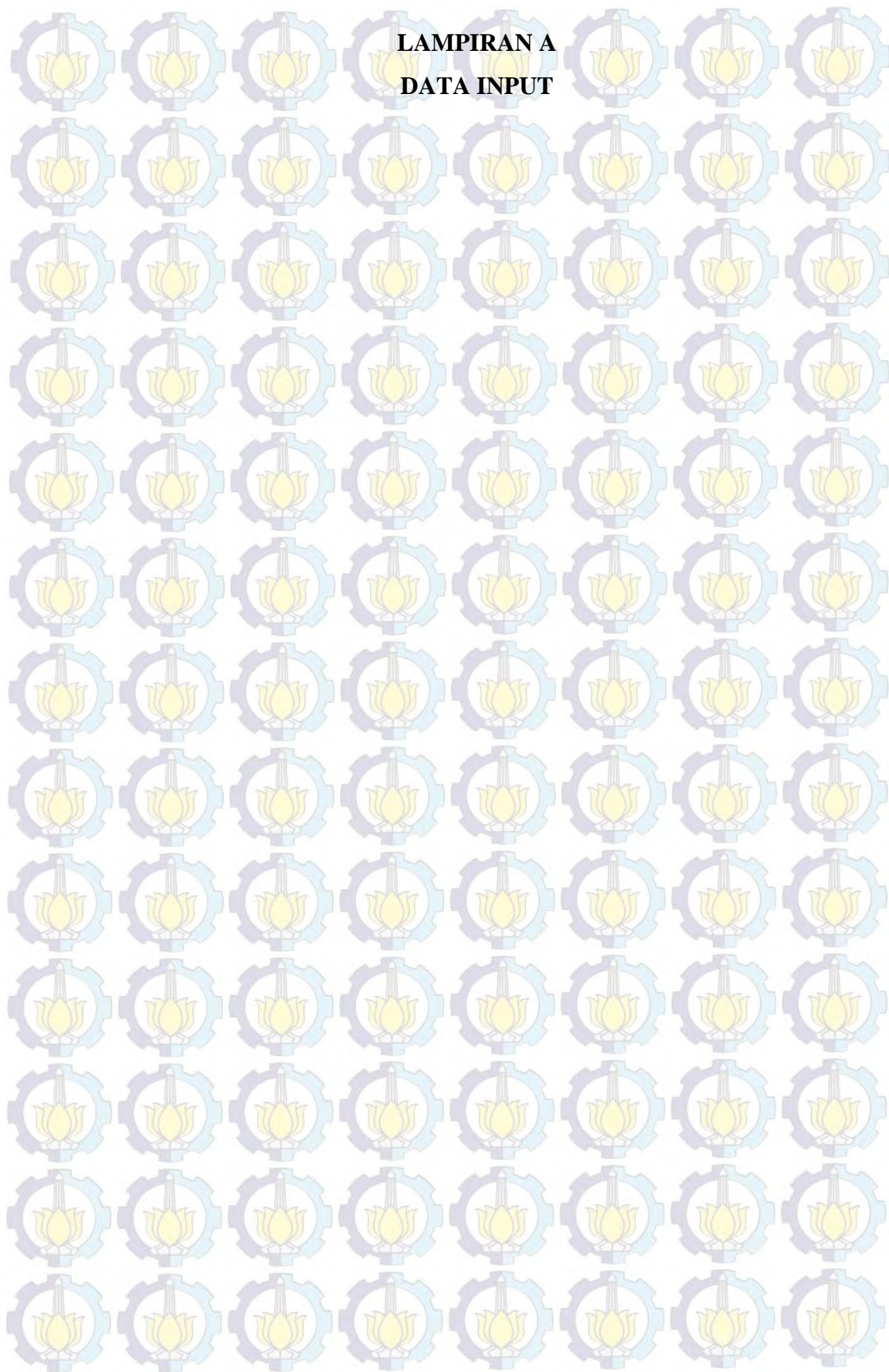
Han, J. and Kamber, M., (2006) *Data Mining: Concepts and Techniques*, Elsevier.

Sunjana. Klasifikasi Data Nasabah sebuah Asuransi Menggunakan Algoritma C4.5 Seminar Nasional Aplikasi Teknologi Informasi 2010 (SNATI 2010). Yogyakarta.

MySQL – The world's most popular open source database,
<http://www.mysql.com/>

RapidMiner, <http://rapid-i.com/content/view/181/190/>

Witten, Ian H. "Text Mining"



NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9109201301	Universitas terbuka	Statistika	3.6	54.94	553	72	80	69.88	3.45	3.59	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109201507	Universitas Wijaya	Manajemen	3.28	66.93	430	80	80	70.28	3.09	3.15	MI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9109201510	ITS Surabaya	Teknik Kelautan	3.46	82.91	477	80	77	76.27	3.27	3.14	MI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9109201508	ITB	Teknik Kimia	3.64	98.9	480	84	81	83.27	3.45	3.54	MI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9109205301	STIKI Malang	Teknik Informatika	3.56	59.94	393	59	78	60.68	3.5	3.5	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205327	Univ. Petra Surabaya	Teknik Informatika	3.71	80.91	550	60	76	73.07	3.6	3.75	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205503	ITS Surabaya	Teknologi Informasi	3.42	58	420	58	78	60.4	3.3	3.67	MTI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9109201306	ITS Surabaya	Teknik Elektro	3.57	92.9	487	92	78	83.47	3.36	3.68	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109201313	ITS Surabaya	Teknik Industri	3.51	100	547	62	78	79.6	3.41	3.64	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205502	ITS Surabaya	Teknik Fisika	3.32	61	407	56	78	60.3	3.3	3.13	MTI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9109205403	ITS Surabaya	Sistem Informasi	3.72	65	517	83	80	74.16	3.8	3.86	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205405	Stikom Surabaya	Sistem Informasi	3.92	82	513	83	75	78.06	3.8	3.92	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205404	Univ. Tarumanegara	Teknik Informatika	3.62	85.71	473	52	85	70.11	3.7	3.75	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205431	Univ. Andalas	Teknik Industri	3.57	73	467	75	85	72.8	3.3	3.75	MP	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205302	Univ. Surabaya	Teknik Informatika	3.89	70	547	65	74	70.7	3.8	3.9	MTI	Profesional	BPPS	Jauh	Tepat Waktu
9109201302	Unika Widya	Teknik Kimia	3.33	66.93	400	64	72	63.28	2.55	3.15	MI	Profesional	Biaya	Jauh	Tepat

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
	Mandala												Sendiri		Waktu
9109201312	STTS	Teknik dan Manajemen Industri	3.4	44.96	410	85	72	63.19	2.95	3.31	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205320	Univ. Petra Surabaya	Sastra Inggris	3.42	61.94	517	46	75	61.18	2.9	3.57	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205318	STAN	Akuntansi	3.58	85.91	487	46	90	69.97	3.4	3.5	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205322	Unair Surabaya	Fisika	3.58	76	430	43	75	60.9	3.3	3.57	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205506	Univ. Islam Indonesia	Teknik Elektro	3.85	87	537	56	85	74.7	3.8	3.79	MTI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9109201314	ITS Surabaya	Teknik Industri	3.53	87.91	443	84	81	78.57	3.41	3.41	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109201505	UPN Veteran Jawa Timur	Teknik Industri	3.22	61.94	380	74	74	64.78	3.05	2.88	MI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9109205311	STIE Pancabhakti Palu	Akuntansi	3.81	94.9	447	54	85	72.67	3.5	3.8	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205305	Univ. Brawijaya Malang	Administrasi Bisnis	3.39	69.93	483	44	74	61.18	3	3.57	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205325	STIKI Malang	Manajemen Informatika & Komputer	3.25	60	420	60	75	61	2.9	3.43	MTI	Profesional	BPPS	Jauh	Tepat Waktu
9109205321	ITB	Teknik Sipil	3.5	84.91	483	48	73	66.67	3.2	3.5	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109205505	ITATS Surabaya	Teknik Informatika	3.53	59	443	58	85	62.9	3.5	3.83	MTI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9109205507	STTS	Teknik Informatika	3.51	64	530	54	85	66.8	3.5	3.63	MTI	Eksekutif	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109201409	Univ. Brawijaya Malang	Teknologi Industri Pertanian	3.39	66	453	65	80	66.5	3.27	3.27	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205401	ITS Surabaya	Teknik Informatika	3.67	77	490	77	80	74.84	3.3	3.86	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205408	Univ. Surabaya	Teknik Informatika	3.56	83	480	74	78	74.99	3.6	3.5	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205415	Univ. Padjadjaran	Teknik Informatika	3.5	74	397	77	78	70.34	3.3	3.79	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205416	Stikom Surabaya	Sistem Informasi	3.69	77	417	77	70	70.04	3.5	3.79	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205417	Univ. Budi Luhur Jakarta	Teknik Informatika	3.78	68	607	69	75	74.37	3.5	3.86	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205413	Univ. Sriwijaya Palembang	Teknik Elektro	3.69	59	477	63	75	63.56	3.6	3.86	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109205418	Stikom Surabaya	Sistem Informasi	3.53	61	493	66	75	65.81	3.6	3.79	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205410	ITS Surabaya	Teknik Informatika	3.64	65	517	60	78	66.9	3.3	3.79	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205419	Univ. Negeri Malang	Fisika	3.56	80	487	77	70	73.54	3.3	3.75	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109207501	Sekolah Tinggi Ilmu Pelayaran	Nautika	3.58	71	467	70	80	69.7	3.5	3.41	MBM	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109205329	Univ. Surabaya	Teknik Informatika	3.13	51.94	397	65	78	60.08	2.4	3.29	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9110201507	ITS Surabaya	Teknik Perkapalan	3.4	61	477	80	80	70.3	3.45	3.12	MI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9110201505	ITS Surabaya	Teknik Sipil	3.85	76	403	75	79	70.7	3.82	3.81	MI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9110201504	Unair Surabaya	Farmasi	3.64	82	470	78	80	75.6	3.68	3.35	MI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9109201511	ITPS	Teknik Mesin	3.25	61.94	367	76	78	65.78	3.27	2.91	MI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109205414	Univ. Kristen Immanuel Yogyakarta	Teknik Informatika	3.44	66.8	400	53	77	60.94	3.3	3.79	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109201406	ITB	Teknik Kimia	3.75	79	530	86	78	79.5	3.68	3.63	MI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9109201408	ITS Surabaya	Teknik Elektro	2.97	78	493	55	73	67.3	2.73	2.91	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109201410	ITS Surabaya	Teknik Kimia	3.75	89	480	76	75	76.7	3.59	3.86	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109201415	Universitas Trunojoyo	Teknik Industri	3.5	62	547	86	79	75.6	3.05	3.68	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201309	Univ. Petra Surabaya	Teknik Mesin	3.82	87	437	76	79	75.1	3.73	3.81	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9109201309	Univ. Sriwijaya Palembang	Teknik Elektro	3.23	70.93	447	88	79	74.48	2.73	2.92	MI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109205308	Institut Informatika Indonesia	Teknik Informatika	3.17	63.94	437	52	77	60.58	2.9	3.5	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109205312	STAN	Akuntansi	3.51	82.91	477	58	70	68.27	3.1	3.65	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9109205324	ITS Surabaya	Teknik Informatika	3.35	58.5	430	57	77	60.25	2.9	3.65	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110205304	Univ. Diponegoro	Teknik Elektro	3.67	58	443	65	80	63.7	3.5	3.9	MTI	Profesional	BPPS	Jauh	Tidak Tepat Waktu
9110205301	Universitas Mercu Buana	Teknik Elektro	3.51	67	450	47	76	60.6	3.5	3.65	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9110205303	Poltek ITS	Teknik Telekomunikasi	3.46	44	380	72	80	60	3.4	3.75	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110205307	ITS Surabaya	Teknik Sipil	3.56	52	380	66	78	60.2	3.4	3.8	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110205314	Univ. Surabaya	Teknik Informatika	3.71	73	430	71	68	67	3.7	3.65	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109202307	UNTAG	Teknik Sipil	3.32	42	407	74	80	60.4	3.9	3.77	MP	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110205310	Pens-ITSPoliteknik Elektronika Negeri Surabaya	Teknologi Informasi	3.46	55	407	61	78	60	3.4	3.65	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201712	ITS Surabaya	Teknik Mesin	3.68	84	490	75	75	75.3	3.18	3.5	MI	Eksekutif	kerjasama	Jauh	Tepat Waktu
9109205501	Univ. Islam Indonesia	Teknik Elektro	3.36	55	353	60	85	60.1	3.2	3.42	MTI	Eksekutif	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109205504	Unika Widya Mandala	Teknik Elektro	3.42	76.93	420	42	75	60.68	3.55	3	MTI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110201412	Pens-ITSPoliteknik	Teknik Telekomunikasi	3.63	73	427	47	70	60	3.32	3.67	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
	Elektronika Negeri Surabaya														
9110201304	ITS Surabaya	Statistika	3.53	38	400	82	79	61.4	3.41	3.32	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9110201314	ITS Surabaya	Teknik Lingkungan	3.51	83	477	65	76	71.6	3.41	3.67	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201305	Univ. Surabaya	Teknik Industri	3.65	83	517	77	79	77.6	3.73	3.65	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201405	UPN Veteran Jawa Timur	Teknik Industri	3	59	447	65	85	65.2	2.73	2.67	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201308	Unair Surabaya	Manajemen	3.33	69	467	74	81	70.5	2.64	3	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9110201401	ITS Surabaya	Teknik Industri	3.81	85	460	76	80	75.7	3.82	3.73	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9110201302	UGM Yogyakarta	Pengolahan Hasil Pertanian	4	88	540	90	84	85.2	4	4	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9110201301	Univ. Surabaya	Teknik Industri	3.76	64	443	77	80	69.1	3.59	3.8	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109201503	ITS Surabaya	Teknik Mesin	3.1	79.91	477	88	77	77.77	3.18	3.06	MI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109201403	ITS Surabaya	Teknik Mesin	3.51	76	480	61	77	68.7	3.41	3.31	MI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109205315	STAN	Akuntansi	3.72	81	477	50	85	68.3	3.1	3.25	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109201402	Univ. Diponegoro	Teknik Industri	3.53	79	460	66	74	69.7	3.59	3.35	MI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9109201308	Univ. Sumatra Utara	Teknik Industri	3.43	99.9	440	76	80	79.37	3.45	3.27	MI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109201303	Stikom Surabaya	Teknik Informatika	3.28	54.94	397	74	70	62.08	2.45	3	MI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110201701	ITS Surabaya	Teknik Elektro	3.97	92	507	75	90	81.5	4	3.89	MI	Eksekutif	kerjasama	Jauh	Tepat Waktu
9110201702	Sekolah Tinggi Teknologi Nurul Jadid	Teknik Elektro	3.53	94	470	65	80	75.3	3.5	3.61	MI	Eksekutif	kerjasama	Dekat	Tepat Waktu
9109201504	ITS Surabaya	Teknik Industri	3.35	75.93	393	82	79	72.58	3.27	3.14	MI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110201707	UGM Yogyakarta	Teknik Kimia	3.86	93	490	75	80	79	4	3.89	MI	Eksekutif	kerjasama	Jauh	Tepat Waktu
9110201705	UGM Yogyakarta	Teknik Geodesi	3.82	71	470	76	75	70.7	4	3.72	MI	Eksekutif	kerjasama	Jauh	Tidak Tepat Waktu
9110201509	ITS Surabaya	Teknik Elektro	3.65	58	387	74	90	66.8	3.64	3.61	MI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9109202404	ITATS Surabaya	Teknik Informatika	3.57	60	440	78	75	67	3.6	3.65	MP	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109201404	Poltek Elka ITS Sby	Teknologi Informasi	3.14	66	453	69	74	66.5	2.45	2.59	MI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110201508	Univ. Surabaya	Manajemen	3.67	70	523	80	65	72	3.5	3.61	MI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9110201510	ITS Surabaya	Teknik Fisika	3.63	63	430	77	75	67.2	3.68	3.33	MI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9109201407	Univ. Negeri Malang	Matematika	3.08	64	417	66	70	62.8	2.55	2.82	MI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109201411	Univ. Brawijaya Malang	Teknik Mesin	3.53	81	500	84	77	78.1	3.45	3.53	MI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9111201532	Univ. Brawijaya Malang	Teknik Elektro	3.64	87	520	74	80	78.3	3.82	3.46	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110205305	STT Telkom Bandung	Teknik Elektro	3.56	64	467	49	75	60.3	3.6	3.6	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109205326	Sekolah Tinggi Teknologi Nurul Jadid	Teknik Informatika	3.21	65	370	59	70	60.2	2.4	3.35	MTI	Profesional	BPPS	Jauh	Tepat Waktu
9109205304	ITS Surabaya	Teknik Informatika	3.34	58.94	457	51	80	60.18	3.1	3.42	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109203302	ITS Surabaya	Teknik Lingkungan	3.46	81.91	403	65	80	69.67	3.59	3.18	MTK	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110205312	Univ. Surabaya	Teknik Informatika	3.53	54	397	65	75	60.1	3.5	3.5	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110205315	ITS Surabaya	Teknik Elektro	3.57	81	573	50	63	68.9	3.5	3.75	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110201413	Pens-ITSPoliteknik Elektronika Negeri	Teknik Telekomunikasi	3.36	58	477	50	78	60	3.55	2.77	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
	Surabaya														
9109207301	ITS Surabaya	Sistem Perkapalan	3.4	63	437	75	80	67.8	3.41	3.18	MBM	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110201417	UPN Veteran Jawa Timur	Teknik Kimia	3.33	64	410	74	70	65.2	3.18	3	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201410	Univ. Petra Surabaya	Arsitektur	3.54	97	523	76	75	80.9	3.59	3.45	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201409	Univ. Petra Surabaya	Teknik Industri	3.66	85	463	70	75	72.9	3.64	3.82	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201414	ITS Surabaya	Teknik Elektro	3.74	96	487	74	75	78.4	3.64	3.68	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110201407	Unair Surabaya	Ekonomi Manajemen	3.14	76	420	50	75	62.8	3.14	2.28	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109201502	UGM Yogyakarta	Teknik Sipil	3.19	86.91	507	80	75	78.47	3.14	1.77	MI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109201501	Universitas Wijaya	Manajemen	3.25	87.91	497	72	78	76.57	3.09	2.69	MI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110202409	ITS Surabaya	Teknik Sipil	3.72	65	420	64	80	64.7	3.7	3.77	MP	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9110202406	ITS Surabaya	Arsitektur	3.63	85	467	53	75	67.8	3.6	3.59	MP	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110202305	Universitas Negeri Jember	Teknik Sipil	3.66	56	413	56	85	60.4	3.6	3.5	MP	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110202303	ITS Surabaya	Teknik Mesin	3.48	89	550	71	80	79.6	3.4	3.41	MP	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109202402	ITB	Bisnis dan	3.36	70	457	68	80	68.6	3	3.73	MP	Profesional	Biaya	Dekat	Tidak

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
		Manajemen											Sendiri		Tepat Waktu
9109201414	ITS Surabaya	Teknik Elektro	3.34	76	433	62	74	66.4	3.05	3.37	MI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110202709	UGM Yogyakarta	Teknik Sipil	3.61	58	377	60	84	61.2	3.59	2.32	MP	Eksekutif	kerjasama	Jauh	Tidak Tepat Waktu
9110201418	ITS Surabaya	Desain Produk Industri	3.4	77	460	75	75	72	2.73	2.86	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9110202701	UK Parahyangan	Teknik Sipil	3.73	81	550	65	82	75.8	3.92	2.32	MP	Eksekutif	kerjasama	Jauh	Tidak Tepat Waktu
9111202703	Univ. Islam Indonesia	Akuntansi	3.72	76	527	80	87	78.4	3.67	2.26	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9109205314	STAN	Akuntansi	3.54	77.91	487	56	75	67.57	3.2	3.75	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9111201318	ITS Surabaya	Teknik Elektro	3.72	93	497	75	95	82.4	3.82	3.73	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205323	Univ. Petra Surabaya	Teknik Informatika	3.6	73.93	443	48	76	62.58	3.2	3.68	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109202304	ITS Surabaya	Teknik Perkapalan	3.43	76	393	64	80	67.4	3.2	3.05	MP	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110205403	Univ. Narotama	Sistem Informasi	3.47	54	433	60	90	62.4	3.1	3	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9111201305	Politeknik Elektronika Negeri	Teknik Elektro	3.21	63	427	69	75	64.6	3.09	2.68	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
	Surabaya														
9111201304	ITS Surabaya	Teknik Mesin	3.32	62	477	62	86	66.4	3.36	3.05	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9111201309	ITS Surabaya	Teknik Mesin	3.61	66	423	74	88	69.6	3.59	3.39	MI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9111201303	Universitas Negeri Surabaya	Kimia	3.29	65	427	70	84	67.3	3.18	2.77	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9111201312	Pens-ITSPoliteknik Elektronika Negeri Surabaya	Teknik Elektro	3.54	72	443	71	82	70.1	3.68	3.18	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205317	STAN	Akuntansi	3.61	82	483	58	85	71.2	3.3	3.5	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9111201315	Univ. Surabaya	Teknik Kimia	3.76	87	457	60	84	72.1	3.82	3.61	MI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9111201306	ITS Surabaya	Teknik Elektro	3.32	65	427	69	78	65.8	2.73	2.77	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9109205309	Unair Surabaya	Akuntansi	3.67	95	547	56	85	77.7	3.4	3.65	MTI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9109201401	Universitas Mercu Buana	Teknik Mesin	3.17	60	427	60	74	60.8	2.45	2.79	MI	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9111201302	ITS Surabaya	Teknik Mesin	3.61	79	490	76	91	77.3	3.41	3.32	MI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9111202303	ITS Surabaya	Arsitektur	3.53	88	437	62	81	71.6	3.7	3.5	MP	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9111205305	Stikom	Sistem	3.71	69	463	62	87	68.1	3.7	3.65	MTI	Profesional	Biaya	Dekat	Tepat

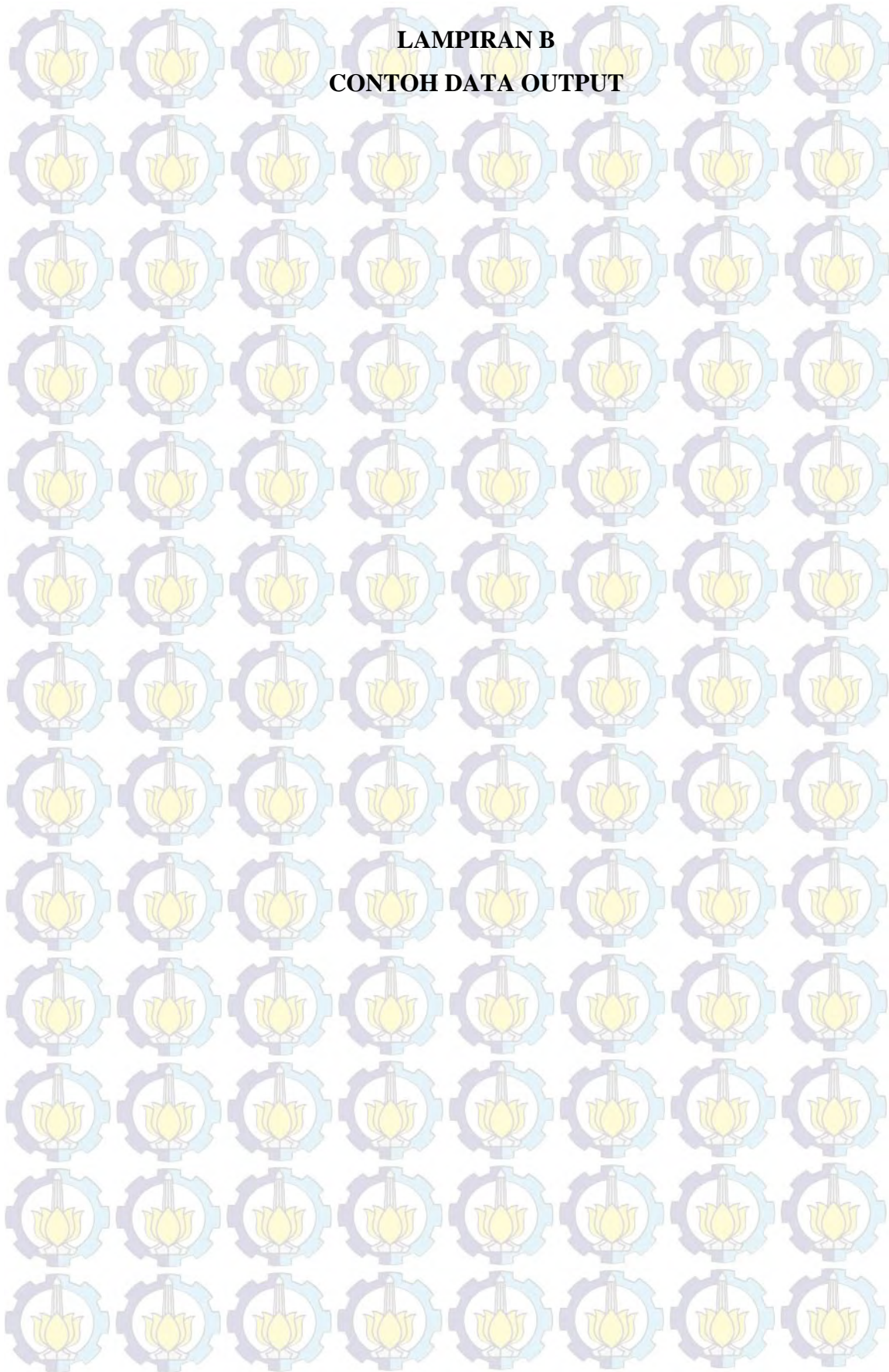
NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
	Surabaya	Informasi											Sendiri		Waktu
9111202714	ITB	Teknik Perminyakan	3.72	80	447	88	84	78.2	3.75	3.88	MP	Profesional	kerjasama	Dekat	Tepat Waktu
9111205322	Pens-ITSPoliteknik Elektronika Negeri Surabaya	Teknik Informatika	3.44	79	513	37	85	65.4	2.7	3.3	MTI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9111205325	Univ. Brawijaya Malang	Teknik Elektro	3.61	74	503	31	82	60.7	3.7	3.5	MTI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9111202701	UGM Yogyakarta	Teknik Elektro	3.81	76	547	80	88	79.8	3.63	3.88	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9110201709	UGM Yogyakarta	Teknik Kimia	3.75	80	500	75	75	74.7	3.78	3.67	MI	Eksekutif	kerjasama	Dekat	Tidak Tepat Waktu
9110202401	ITS Surabaya	Teknik Mesin	3.85	81	467	100	90	83.7	4	4	MP	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110202307	ITS Surabaya	Arsitektur	3.82	74	477	65	85	70.7	3.9	3.77	MP	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9110202705	Univ. Brawijaya Malang	Teknik Pengairan	3.46	55	493	75	83	68.4	3.67	2.09	MP	Eksekutif	kerjasama	Jauh	Tidak Tepat Waktu
9111202702	Univ. Indonesia	Teknik Elektro	3.78	74	517	81	88	77.9	3.75	3.88	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9110202703	UK Parahyangan	Teknik Sipil	3.63	77	533	54	84	70.7	3.67	2.26	MP	Eksekutif	kerjasama	Jauh	Tidak Tepat Waktu
9111201307	ITS Surabaya	Teknik	3.61	73	516	70	86	73.7	2.91	3.27	MI	Eksekutif	Biaya	Dekat	Tepat

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
		Perkapalan											Sendiri		Waktu
9110205411	Pens-ITSPoliteknik Elektronika Negeri Surabaya	Teknik Elektro	3.42	73	437	52	80	63.9	3.2	3.5	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9111201412	ITS Surabaya	Teknik Elektro	3.38	66	477	60	81	66	3.18	3.28	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9111201414	Pens-ITSPoliteknik Elektronika Negeri Surabaya	Teknik Elektro	3.49	73	457	81	80	73.4	3.36	3.17	MI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9111202401	ITS Surabaya	Arsitektur	3.43	60	480	54	78	62	3.6	3.59	MP	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9111202704	ITS Surabaya	Teknik Mesin	3.72	80	517	70	85	75.8	3.63	3.75	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9111202706	UGM Yogyakarta	Teknik Kimia	3.86	88	490	76	86	79	4	3.88	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9111202719	ITB	Teknik Mesin	3.81	61	493	73	87	70.4	3.75	3.63	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9111202721	ITS Surabaya	Teknik Mesin	3.64	75	447	77	85	73.6	3.63	3.63	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9111202715	Univ. Sumatra Utara	Teknik Elektro	3.53	74	437	74	75	69.8	3.5	3.5	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9111202718	Swiss Germany University	Mechatronics	3.56	66	553	79	78	74.9	3.38	3.63	MP	Profesional	kerjasama	Dekat	Tepat Waktu
9111202722	ITS Surabaya	Teknik Elektro	3.78	84	473	73	86	76.1	3.75	3.63	MP	Profesional	kerjasama	Jauh	Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9111202712	UGM Yogyakarta	Teknik Elektro	3.56	79	527	78	85	78.3	3.63	3.63	MP	Profesional	kerjasama	Jauh	Tepat Waktu
9111205311	Pens-ITSPoliteknik Elektronika Negeri Surabaya	Teknologi Informasi	3.54	77	410	58	86.3	67.56	3.5	3.55	MTI	Profesional	Biaya Sendiri	Dekat	Tepat Waktu
9111205315	ITS Surabaya	Teknik Informatika	3.51	72	503	40	84	63.2	3.7	3.45	MTI	Eksekutif	Biaya Sendiri	Dekat	Tepat Waktu
9111205302	ITS Surabaya	Teknik Mesin	3.68	62	480	64	72	64.4	3.8	3.55	MTI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9111205306	Stikom Surabaya	Manajemen Informatika & Komputer	3.63	69	487	91	89	78.2	3.6	3.65	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9111205307	ITS Surabaya	Teknik Informatika	3.76	100	523	81	93	86.9	3.8	3.75	MTI	Profesional	Biaya Sendiri	Jauh	Tepat Waktu
9111205303	Univ. Brawijaya Malang	Teknik Mesin	3.86	84	500	69	81	75.3	3.9	3.9	MTI	Eksekutif	Biaya Sendiri	Jauh	Tepat Waktu
9110205311	Stikom Surabaya	Sistem Informasi	3.42	75	543	69	70	72.4	3.6	3.3	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110201503	Universitas Islam Indonesia	Teknik Elektro	3.32	64	393	77	79	67.5	3.18	3.06	MI	Eksekutif	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9111201319	Unika Widya Mandala	Teknik Kimia	3.63	89	490	65	92	77.2	3.64	3.39	MI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9111205324	Univ. Udayana	Teknik Elektro	3.69	70	493	90	84.7	77.74	3.8	3.4	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu

NRP	Nama PT S1	Jurusan S1	IPK S1	GMAT	TOEFL	M.Bid	Wwn cara	Skor Akhir	IPS S 1	IPS S2	Prog. Studi	Bidang Keahlian	Biaya	Jarak S1-S2	CLASS
9110201704	Univ. Indonesia	Marketing Manajemen	3.44	78	499	71	80	73.3	3.5	3.5	MI	Eksekutif	kerjasama	Dekat	Tidak Tepat Waktu
9109203301	Univ. Islam Indonesia	Teknik Lingkungan	3.32	67.93	420	50	78	60.98	3.18	3.32	MTK	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9109207401	ITS Surabaya	Teknik Sipil	3.35	67	453	85	75	71.8	3.32	2	MBM	Profesional	Biaya Sendiri	Jauh	Tidak Tepat Waktu
9111202713	Univ. Diponegoro	Teknik Mesin	3.47	66	457	80	83	71.6	3.25	3.5	MP	Profesional	kerjasama	Jauh	Tidak Tepat Waktu
9111202708	Univ. Jendral A.Yani	Teknik Mesin	3.72	67	440	73	82	69	3.63	3.63	MP	Profesional	kerjasama	Jauh	Tidak Tepat Waktu
9111202716	Univ. Bung Hatta	Teknik Elektro	3.67	78	430	71	88	72.5	4	3.63	MP	Profesional	kerjasama	Dekat	Tidak Tepat Waktu
9111202707	ITB	Teknik Mesin	3.61	84	543	65	84	76.7	3.75	3.75	MP	Profesional	kerjasama	Jauh	Tidak Tepat Waktu
9111205318	Stikom Surabaya	Sistem Informasi	3.63	78	480	34	84	62.6	3.7	3.55	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu
9110205410	Pens-ITSPoliteknik Elektronika Negeri Surabaya	Teknik Informatika	3.58	65	507	46	75	61.7	3.5	3.7	MTI	Profesional	Biaya Sendiri	Dekat	Tidak Tepat Waktu

LAMPIRAN B
CONTOH DATA OUTPUT



Pohon Keputusan:

bidang_keahlian = Eksekutif (Tidak Tepat Waktu = 3, Tepat Waktu = 0) : **Tidak Tepat Waktu**

bidang_keahlian = Profesional (Tidak Tepat Waktu = 12, Tepat Waktu = 15) : ?

biaya = Biaya Sendiri (Tidak Tepat Waktu = 8, Tepat Waktu = 4) : ?

status_program_studi = Sebidang (Tidak Tepat Waktu = 5, Tepat Waktu = 4) : ?

wawancara_huruf = B (Tidak Tepat Waktu = 2, Tepat Waktu = 0) : **Tidak Tepat Waktu**

wawancara_huruf = C (Tidak Tepat Waktu = 3, Tepat Waktu = 3) : ?

ipk_s1_huruf = A (Tidak Tepat Waktu = 0, Tepat Waktu = 0) :

ipk_s1_huruf = B (Tidak Tepat Waktu = 0, Tepat Waktu = 1) : **Tepat Waktu**

ipk_s1_huruf = C (Tidak Tepat Waktu = 1, Tepat Waktu = 2) : ?

ips_s2_huruf = A (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

ips_s2_huruf = B (Tidak Tepat Waktu = 0, Tepat Waktu = 2) : **Tepat Waktu**

ips_s2_huruf = C (Tidak Tepat Waktu = 1, Tepat Waktu = 0) : **Tidak Tepat Waktu**

ips_s2_huruf = E (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

ipk_s1_huruf = D (Tidak Tepat Waktu = 2, Tepat Waktu = 0) : **Tidak Tepat Waktu**

wawancara_huruf = D (Tidak Tepat Waktu = 0, Tepat Waktu = 1) : **Tepat Waktu**

status_program_studi = tidak sebidang (Tidak Tepat Waktu = 3, Tepat Waktu = 0) : **Tidak**

Tepat Waktu

biaya = kerjasama (Tidak Tepat Waktu = 4, Tepat Waktu = 11) : ?

m_bid_huruf = A (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

m_bid_huruf = B (Tidak Tepat Waktu = 2, Tepat Waktu = 10) : ?

akreditasi_s1 = A (Tidak Tepat Waktu = 0, Tepat Waktu = 8) : **Tepat Waktu**

akreditasi_s1 = B (Tidak Tepat Waktu = 1, Tepat Waktu = 1) : ?

toefl_huruf = B (Tidak Tepat Waktu = 0, Tepat Waktu = 1) : **Tepat Waktu**

toefl_huruf = D (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

toefl_huruf = C (Tidak Tepat Waktu = 1, Tepat Waktu = 0) : **Tidak Tepat Waktu**

toefl_huruf = E (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

akreditasi_s1 = C (Tidak Tepat Waktu = 1, Tepat Waktu = 1) : ?

gmat_huruf = A (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

gmat_huruf = B (Tidak Tepat Waktu = 0, Tepat Waktu = 1) : **Tepat Waktu**

gmat_huruf = C (Tidak Tepat Waktu = 1, Tepat Waktu = 0) : **Tidak Tepat Waktu**

gmat_huruf = D (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

gmat_huruf = E (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

m_bid_huruf = C (Tidak Tepat Waktu = 2, Tepat Waktu = 1) : ?

ips_s1_huruf = A (Tidak Tepat Waktu = 2, Tepat Waktu = 0) : **Tidak Tepat Waktu**

ips_s1_huruf = B (Tidak Tepat Waktu = 0, Tepat Waktu = 1) : **Tepat Waktu**

ips_s1_huruf = C (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tidak Tepat Waktu**

ips_s1_huruf = D (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tidak Tepat Waktu**

m_bid_huruf = D (Tidak Tepat Waktu = 0, Tepat Waktu = 0) : **Tepat Waktu**

Rule:

1. if (bidang_keahlian == Eksekutif) then **Tidak Tepat Waktu** (id = 1)

2. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == B) then **Tidak Tepat Waktu** (id = 5)

3. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == C AND ipk_s1_huruf == A) then (id = 7)

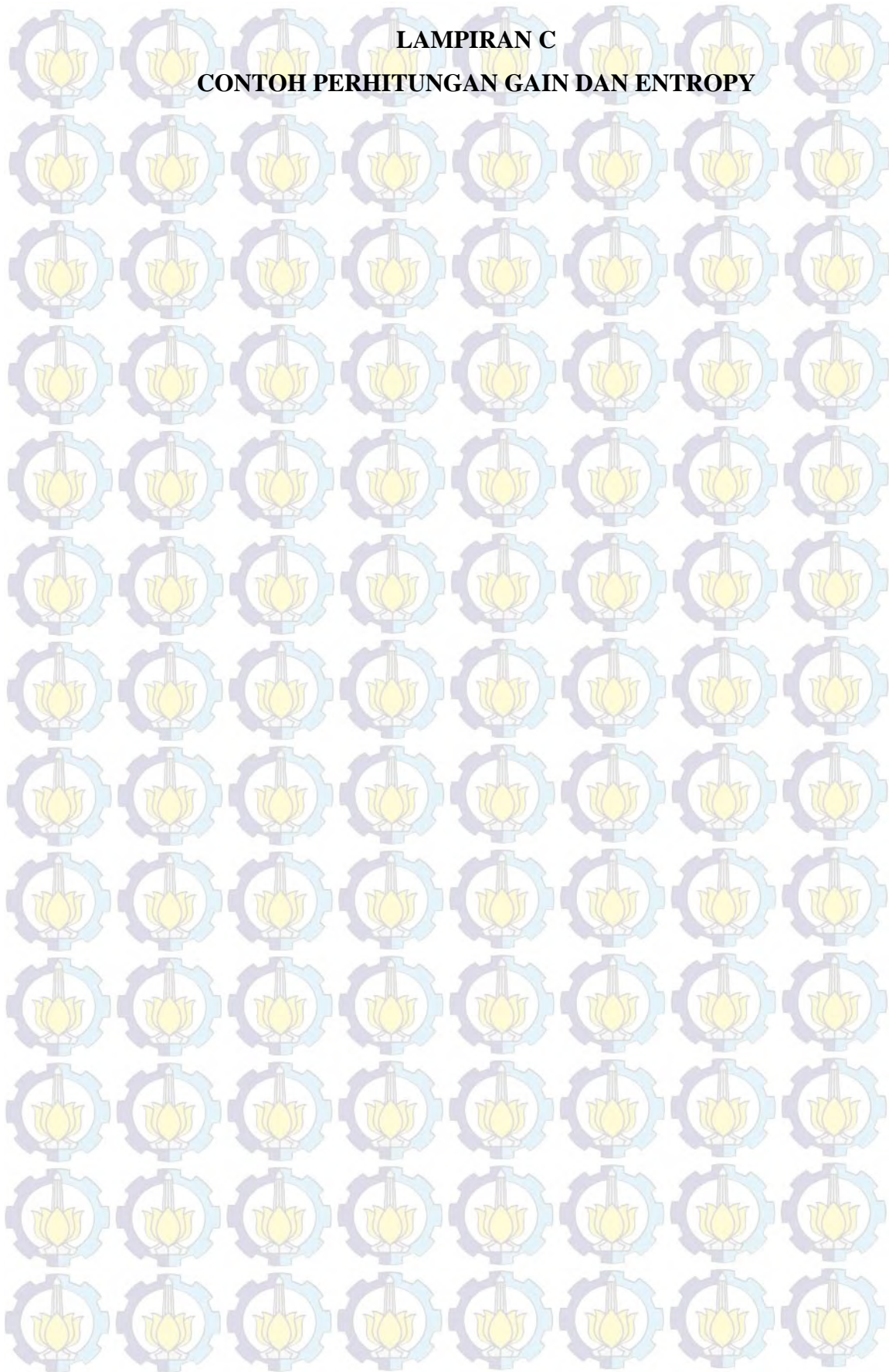
4. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == C AND ipk_s1_huruf == B) then **Tepat Waktu** (id = 8)

5. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == C AND ipk_s1_huruf == C AND ips_s2_huruf == A) then **Tepat Waktu** (id = 10)

6. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == C AND ipk_s1_huruf == C AND ips_s2_huruf ==

- B) then **Tepat** Waktu (id = 11)
7. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == C AND ipk_s1_huruf == C AND ips_s2_huruf == C) then **Tidak** **Tepat** Waktu (id = 12)
8. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == C AND ipk_s1_huruf == C AND ips_s2_huruf == E) then **Tepat** Waktu (id = 13)
9. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == C AND ipk_s1_huruf == D) then **Tidak** **Tepat** Waktu (id = 14)
10. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == Sebidang AND wawancara_huruf == D) then **Tepat** Waktu (id = 15)
11. if (bidang_keahlian == Profesional AND biaya == Biaya Sendiri AND status_program_studi == tidak sebidang) then **Tidak** **Tepat** Waktu (id = 16)
12. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == A) then **Tepat** Waktu (id = 18)
13. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == A) then **Tepat** Waktu (id = 20)
14. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == B A ND toefl_huruf == B) then **Tepat** Waktu (id = 22)
15. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == B A ND toefl_huruf == D) then **Tepat** Waktu (id = 23)
16. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == B A ND toefl_huruf == C) then **Tidak** **Tepat** Waktu (id = 24)
17. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == B A ND toefl_huruf == E) then **Tepat** Waktu (id = 25)
18. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == C AND gmat_huruf == A) then **Tepat** Waktu (id = 27)
19. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == C AND gmat_huruf == B) then **Tepat** Waktu (id = 28)
20. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == C AND gmat_huruf == C) then **Tidak** **Tepat** Waktu (id = 29)
21. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == C AND gmat_huruf == D) then **Tepat** Waktu (id = 30)
22. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == B AND akreditasi_s1 == C AND gmat_huruf == E) then **Tepat** Waktu (id = 31)
23. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == C AND ips_s1_huruf == A) then **Tidak** **Tepat** Waktu (id = 33)
24. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == C AND ips_s1_huruf == B) then **Tepat** Waktu (id = 34)
25. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == C AND ips_s1_huruf == C) then **Tidak** **Tepat** Waktu (id = 35)
26. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == C AND ips_s1_huruf == D) then **Tidak** **Tepat** Waktu (id = 36)
27. if (bidang_keahlian == Profesional AND biaya == kerjasama AND m_bid_huruf == D) then **Tepat** Waktu (id = 37)

LAMPIRAN C
CONTOH PERHITUNGAN GAIN DAN ENTROPY



C45 » Perhitungan C4.5

Opsi: Hapus Semua Data

NO	ATRIBUT GAIN RATIO MAX	ATRIBUT	NILAI ATRIBUT	JUMLAH KASUS TOTAL	JUMLAH KASUS TIDAK TEPAT WAKTU	JUMLAH KASUS TEPAT WAKTU	ENTROPY	INFORMATION GAIN	SPLIT INFO	GAIN RATIO
1	skor_akhir	Total	Total	107	41	66	0.9603			0
2	skor_akhir	akreditasi_s1	A	65	23	42	0.9375	0.0146	1.2772	0.0114
3	skor_akhir	akreditasi_s1	B	32	12	20	0.9544	0.0146	1.2772	0.0114
4	skor_akhir	akreditasi_s1	C	10	6	4	0.971	0.0146	1.2772	0.0114
5	skor_akhir	program_studi	MBM	1	1	0	0	0.0658	1.6587	0.0397
6	skor_akhir	program_studi	MI	45	13	32	0.8673	0.0658	1.6587	0.0397
7	skor_akhir	program_studi	MP	18	10	8	0.9911	0.0658	1.6587	0.0397
8	skor_akhir	program_studi	MTI	41	15	26	0.9474	0.0658	1.6587	0.0397
9	skor_akhir	program_studi	MTK	2	2	0	0	0.0658	1.6587	0.0397
10	skor_akhir	status_program_studi	Sebidang	70	29	41	0.9787	0.0057	0.9303	0.0061

NO	ATRIBUT GAIN RATIO MAX	ATRIBUT	NILAI ATRIBUT	JUML AH KASU S TOTAL	JUML AH KASU S TIDAK TEPAT WAKT U	JUM LAH KAS US TEP AT WA KTU	ENTR OPY	INFORMAT ION GAIN	SPLIT INFO	GAIN RATIO
11	skor_akhir	status_program _studi	tidak sebidang	37	12	25	0.909	0.0057	0.9303	0.0061
12	skor_akhir	bidang_keahlian	Eksekutif	30	14	16	0.996 8	0.0083	0.856	0.0097
13	skor_akhir	bidang_keahlian	Profesio nal	77	27	50	0.934 6	0.0083	0.856	0.0097
14	skor_akhir	biaya	Biaya Sendiri	88	35	53	0.969 6	0.0031	0.7865	0.0039
15	skor_akhir	biaya	BPPS	3	1	2	0.918 3	0.0031	0.7865	0.0039
16	skor_akhir	biaya	kerjasam a	16	5	11	0.896	0.0031	0.7865	0.0039
17	skor_akhir	ipk_s1	A	11	1	10	0.439 5	0.106	2.0506	0.0517
18	skor_akhir	ipk_s1	B	36	9	27	0.811 3	0.106	2.0506	0.0517
19	skor_akhir	ipk_s1	C	37	16	21	0.986 8	0.106	2.0506	0.0517
20	skor_akhir	ipk_s1	D	17	12	5	0.874	0.106	2.0506	0.0517
21	skor_akhir	ipk_s1	E	6	3	3	1	0.106	2.0506	0.0517

NO	ATRIBUT GAIN RATIO MAX	ATRIBUT	NILAI ATRIBUT	JUML AH KASU S TOTAL	JUML AH KASU S TIDAK TEPAT WAKTU	JUMLAH KASUS TEPAT WAKTU	ENTROPY	INFORMATION GAIN	SPLIT INFO	GAIN RATIO
22	skor_akhir	ipk_s2	A	7	1	6	0.5917	0.0672	1.9756	0.034
23	skor_akhir	ipk_s2	B	39	10	29	0.8213	0.0672	1.9756	0.034
24	skor_akhir	ipk_s2	C	36	17	19	0.9978	0.0672	1.9756	0.034
25	skor_akhir	ipk_s2	D	20	9	11	0.9928	0.0672	1.9756	0.034
26	skor_akhir	ipk_s2	E	5	4	1	0.7219	0.0672	1.9756	0.034
27	skor_akhir	gmat	A	16	4	12	0.8113	0.0286	1.7863	0.016
28	skor_akhir	gmat	B	45	17	28	0.9565	0.0286	1.7863	0.016
29	skor_akhir	gmat	C	38	15	23	0.9678	0.0286	1.7863	0.016
30	skor_akhir	gmat	D	7	4	3	0.9852	0.0286	1.7863	0.016
31	skor_akhir	gmat	E	1	1	0	0	0.0286	1.7863	0.016
32	skor_akhir	toefl	A	0	0	0	0	0.0413	1.6102	0.0256

NO	ATRIBUT GAIN RATIO MAX	ATRIBUT	NILAI ATRIBUT	JUML AH KASU S TOTAL	JUML AH KASU S TIDAK TEPAT WAKT U	JUM LAH KAS US TEP AT WA KTU	ENTR OPY	INFORMAT ION GAIN	SPLIT INFO	GAIN RATIO
33	skor_akhir	toefl	B	14	6	8	0.985 2	0.0413	1.6102	0.0256
34	skor_akhir	toefl	C	48	13	35	0.842 7	0.0413	1.6102	0.0256
35	skor_akhir	toefl	D	41	19	22	0.996 1	0.0413	1.6102	0.0256
36	skor_akhir	toefl	E	4	3	1	0.811 3	0.0413	1.6102	0.0256
37	skor_akhir	m_bid	A	7	3	4	0.985 2	0.072	1.8453	0.039



AMELIA HALIM, S.KOM. lahir pada tanggal 16 Mei 1989 di Mojokerto, Jawa Timur, sebagai anak pertama dari dua bersaudara pasangan Maslim Suhardi dan Octavia Vimaladevi. Penulis mulai menempuh pendidikan formal usia 3 tahun di TK Taruna Nusa Harapan Mojokerto. Setelah itu penulis meneruskan studinya di SD Taruna Nusa Harapan dan SMP Taruna Nusa Harapan

Mojokerto. Penulis melanjutkan SMA di Malang yaitu SMAK Kolese Santo Yusup Malang. Setelah lulus dari SMA, Penulis memutuskan untuk melanjutkan perguruan tingginya di perguruan tinggi swasta Surabaya. Penulis memperoleh gelar S.Kom Pada bulan Januari 2012 dari Sekolah Tinggi Teknik Surabaya. Setelah lulus dari perguruan tinggi, penulis menjadi freelance programmer pada suatu perusahaan asing selama 1 tahun 4 bulan. Tak Berselang lama setelah berhenti bekerja sebagai freelance programmer, penulis melamar pekerjaan menjadi marketing di perusahaan Swasta hingga tesis ini dibuat.

DATA PRIBADI PENULIS :

Nama : Amelia Halim, S.Kom.
Alamat : Jalan Letkol Sumarjo No. 96 Mojokerto
Email : Amell_friend@yahoo.com